

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

<p>(51) International Patent Classification ⁷ : C07K 14/705, C12N 15/12, C12Q 1/68, C12N 5/10, C07K 16/28, G01N 33/53, A61K 38/17</p>	A2	<p>(11) International Publication Number: WO 00/40614</p> <p>(43) International Publication Date: 13 July 2000 (13.07.00)</p>									
<p>(21) International Application Number: PCT/US99/29996</p> <p>(22) International Filing Date: 20 December 1999 (20.12.99)</p> <p>(30) Priority Data:</p> <table border="0"> <tr> <td>60/114,220</td> <td>30 December 1998 (30.12.98)</td> <td>US</td> </tr> <tr> <td>60/120,018</td> <td>29 January 1999 (29.01.99)</td> <td>US</td> </tr> <tr> <td>60/140,415</td> <td>22 June 1999 (22.06.99)</td> <td>US</td> </tr> </table> <p>(71) Applicant (for all designated States except US): BETH ISRAEL DEACONESS MEDICAL CENTER, INC. [US/US]; 330 Brookline Avenue, Boston, MA 02215 (US).</p> <p>(72) Inventor; and</p> <p>(75) Inventor/Applicant (for US only): SCHARENBERG, Andrew, M. [US/US]; 12 Skyview Road, Lexington, MA 02420 (US).</p> <p>(74) Agent: PLUMER, Elizabeth, R.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).</p>		60/114,220	30 December 1998 (30.12.98)	US	60/120,018	29 January 1999 (29.01.99)	US	60/140,415	22 June 1999 (22.06.99)	US	<p>(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).</p> <p>Published Without international search report and to be republished upon receipt of that report.</p>
60/114,220	30 December 1998 (30.12.98)	US									
60/120,018	29 January 1999 (29.01.99)	US									
60/140,415	22 June 1999 (22.06.99)	US									
<p>(54) Title: CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY</p> <div data-bbox="315 1127 1239 1465"> </div>											
<p>(57) Abstract</p> <p>Nucleic acids encoding SOC/CRAC calcium channel polypeptides, including fragments and biologically functional variants thereof and encoded polypeptides are provided. The nucleic acids and polypeptides disclosed herein are useful as therapeutic and diagnostic agents. Agents that selectively bind to the foregoing polypeptides and genes also are provided.</p>											

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY**Field of the Invention**

This invention relates to nucleic acids coding for a novel family of calcium channel polypeptides, the encoded polypeptides, unique fragments of the foregoing, and methods of making and using same.

Background of the Invention

Calcium channels are membrane-spanning, multi-subunit proteins that facilitate the controlled transport ("flux") of Ca^{2+} ions into and out of cells. Cells throughout the animal kingdom, and at least some bacterial, fungal and plant cells, possess one or more types of calcium channels. In general, "excitable" cells, such as neurons of the central nervous system, peripheral nerve cells, and muscle cells, including those of skeletal muscles, cardiac muscles, and venous and arterial smooth muscles, possess voltage-dependent calcium channels. In a voltage-dependent calcium channel, the transport of Ca^{2+} ions into and out of the cells requires a certain minimal level of depolarization (the difference in potential between the inside of the cell bearing the channel and the extracellular environment) with the rate of Ca^{2+} cell flux dependent on the difference in potential. In "non-excitable" cells, calcium influx is thought to occur predominantly in response to stimuli which cause the release of calcium from intracellular stores. This process, termed *store operated calcium influx*, is not well understood.

Characterization of a particular type of calcium channel by analysis of whole cells is complicated by the presence of mixed populations of different types of calcium channels in the majority of cells. Although single-channel recording methods can be used to examine individual calcium channels, such analysis does not reveal information related to the molecular structure or biochemical composition of the channel. Furthermore, in this type of analysis, the channel is isolated from other cellular constituents that might be important for the channel's natural functions and pharmacological interactions. To study the calcium channel structure-function relationship, large amounts of pure channel protein are needed. However, acquiring large amounts of pure protein is difficult in view of the complex nature of these multisubunit proteins, the varying concentrations of calcium channel proteins in tissue sources, the presence of mixed populations of calcium channel proteins in tissues, and the modifications of the native protein that can occur during the isolation procedure.

Summary of the Invention

The invention is based on the identification of a novel family of calcium channel polypeptides and the molecular cloning and partial characterization of a novel member of this family that is expressed predominantly in human hematopoietic cells, liver, and kidney. This newly identified family of calcium channel polypeptides is designated, "SOC" or "CRAC" or "ICRAC", for Store Operated Channels or Calcium Release Activated Channels. Although not wishing to be bound to any particular theory or mechanism, it is believed that the SOC/CRAC calcium channel polypeptides are transmembrane polypeptides that modulate Ca^{2+} flux "into" and "out of" a cell, for example, in certain instances they may be activated upon depletion of Ca^{2+} from intracellular calcium stores, allowing Ca^{2+} influx into the cell. Accordingly, the compositions disclosed herein are believed to be useful for modulating calcium transport into and out of such intracellular stores and for the treatment of disorders that are characterized by aberrant calcium transport into and out of such intracellular stores. In particular, we believe that the SOC/CRAC calcium channel polypeptides disclosed herein play an important role in the influx of extracellular calcium by mediating the refilling of intracellular calcium stores following their depletion. Accordingly, we believe that the compositions for expressing functional SOC/CRAC calcium channel polypeptides in cells, as disclosed herein, are useful for treating patients having conditions that are characterized by reduced extracellular calcium influx into their SOC/CRAC-expressing cells. Additionally, the compositions of the invention are useful for delivering therapeutic and/or imaging agents to cells which preferentially express SOC/CRAC calcium channel polypeptides and, in particular, for delivering such agents to hematopoietic cells, liver, heart, spleen, and kidney to modulate proliferation and growth of these cells. Moreover, in view of the importance of cellular calcium levels to cell viability, we believe that SOC-2/CRAC-1, SOC-3/CRAC-2, and SOC-4/CRAC-3 as disclosed herein, and/or other members of the SOC/CRAC family of calcium channel polypeptides, represent an ideal target for designing and/or identifying (e.g., from molecular libraries) small molecule inhibitors that block lymphocyte proliferation, as well as other binding agents that selectively bind to SOC/CRAC polypeptides to which drugs or toxins can be conjugated for delivery to SOC/CRAC polypeptide expressing cells.

The invention is based, in part, on the molecular cloning and sequence analysis of the novel SOC/CRAC calcium channel molecules disclosed herein (also referred to as a "SOC-2/CRAC-1 molecule," a "SOC-3/CRAC-2 molecule," and/or "SOC-4/CRAC-3 molecule") that are predominantly expressed in human hematopoietic cells, liver, spleen, heart, and

kidney (SOC-2/CRAC-1), kidney and colon (SOC-3/CRAC-2), and prostate (SOC-4/CRAC-3 molecule). As used herein, a "SOC/CRAC molecule" embraces a "SOC/CRAC calcium channel nucleic acid" (or "SOC/CRAC nucleic acid") and a "SOC/CRAC calcium channel polypeptide" (or "SOC/CRAC polypeptide"). Homologs and alleles also are embraced within the meaning of a SOC/CRAC calcium channel molecule.

According to one aspect of the invention, isolated SOC/CRAC nucleic acids which code for one or more member(s) of the SOC/CRAC family of calcium channel polypeptides or unique fragments thereof are provided. The isolated nucleic acids refer to one or more of the following:

(a) nucleic acid molecules which hybridize under stringent conditions to a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;

(b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;

(c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and

(d) complements of (a), (b) or (c).

The invention in another aspect provides an isolated nucleic acid molecule selected from the group consisting of (a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:29, and SEQ ID NO:31, (b) complements of (a), provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to any sequence selected from a sequence group consisting of (1) sequences having the SEQ. ID NOS. or GenBank accession numbers of Table I, (2) complements of (1), and (3) fragments of (1) and (2).

According to yet another aspect of the invention, isolated SOC/CRAC polypeptides are provided. The isolated SOC/CRAC polypeptide molecules are encoded by one or more SOC/CRAC nucleic acid molecules of the invention. Preferably, the SOC/CRAC polypeptide contains one or more polypeptides selected from the group consisting of the polypeptides having SEQ. ID Nos. 2, 4, 6, 8, 24, 26, 28, 30, and 32. In other embodiments, the isolated polypeptide may be a fragment or variant of the foregoing SOC/CRAC polypeptide molecules of sufficient length to represent a sequence unique within the human genome, and identifying

with a polypeptide that functions as a calcium channel, provided that the fragment excludes a sequence of contiguous amino acids identified in Table II, and/or excludes a sequence of contiguous amino acids encoded for by a nucleic acid sequence identified in Table I. In another embodiment, immunogenic fragments of the polypeptide molecules described above are provided.

According to another aspect of the invention, isolated SOC/CRAC binding agents (e.g., polypeptides) are provided which selectively bind to a SOC/CRAC molecule (e.g., a SOC/CRAC polypeptide encoded by the isolated nucleic acid molecules of the invention). Preferably, the isolated binding agents selectively bind to a polypeptide which comprises the sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32, or unique fragments thereof. In the preferred embodiments, the isolated binding polypeptides include antibodies and fragments of antibodies (e.g., Fab, F(ab)₂, Fd and antibody fragments which include a CDR3 region which binds selectively to a SOC/CRAC polypeptide). Preferably, the antibodies for human therapeutic applications are human antibodies.

According to another aspect of the invention, a pharmaceutical composition containing a pharmaceutically effective amount of an isolated SOC/CRAC nucleic acid, an isolated SOC/CRAC polypeptide, or an isolated SOC/CRAC binding polypeptide in a pharmaceutically acceptable carrier also is provided. The pharmaceutical compositions are useful in accordance with therapeutic methods disclosed herein.

According to yet another aspect of the invention, a method for isolating a SOC/CRAC molecule is provided. The method involves:

a) contacting a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample that is believed to contain one or more SOC/CRAC molecules, under conditions to form a complex of the SOC/CRAC nucleic acid or the SOC/CRAC binding polypeptide and the SOC/CRAC molecule;

b) detecting the presence of the complex;

c) isolating the SOC/CRAC molecule from the complex; and

d) determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity. As used herein "SOC/CRAC calcium channel activity" refers to the transport of Ca²⁺ into and out of intracellular stores that is mediated by a SOC/CRAC

polypeptide. In general, the SOC/CRAC calcium channel activity is initiated by a reduction or depletion of intracellular calcium stores.

In certain embodiments, the SOC/CRAC nucleic acid is a SOC-2/CRAC-1 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 27, or complements thereof); in certain other
5 embodiments, the SOC/CRAC nucleic acid is a SOC-3/CRAC-2 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 29, or complements thereof); in further embodiments, the SOC/CRAC nucleic acid is a SOC-4/CRAC-3 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 31, or complements thereof). In yet other embodiments, the SOC/CRAC polypeptide is a SOC-2/CRAC-1 binding polypeptide (e.g., an antibody that selectively binds to a SOC-
10 2/CRAC-1 polypeptide). In yet further embodiments, the SOC/CRAC polypeptide is a SOC-3/CRAC-2 binding polypeptide (e.g., an antibody that selectively binds to a SOC-3/CRAC-2 polypeptide). In some embodiments, the SOC/CRAC polypeptide is a SOC-4/CRAC-3 binding polypeptide (e.g., an antibody that selectively binds to a SOC-4/CRAC-3 polypeptide). In the preferred embodiments, the isolated binding polypeptides include
15 antibodies and fragments of antibodies (e.g., Fab, F(ab)₂, Fd and antibody fragments which include a CDR3 region which binds selectively to a SOC-2/CRAC-1, to a SOC-3/CRAC-2, and/or to a SOC-4/CRAC-3 polypeptide). Preferably the isolated binding polypeptides or other binding agents selectively bind to a single SOC/CRAC molecule, i.e., are capable of distinguishing between different members of the SOC/CRAC family. Accordingly, one or
20 more SOC/CRAC binding agents can be contained in a single composition (e.g., a pharmaceutical composition) to identify multiple SOC/CRAC molecules *in vivo* or *in vitro*.

According to yet another aspect of the invention, a method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity is provided. The method involves:

25 a) contacting a SOC/CRAC polypeptide with a candidate agent suspected of modulating SOC/CRAC calcium channel activity, under conditions sufficient to allow the candidate agent to interact selectively with (e.g. bind to) the SOC/CRAC polypeptide;

 b) detecting a Ca²⁺ concentration of step (b) associated with the SOC/CRAC calcium channel activity of the SOC/CRAC polypeptide in the presence of the candidate agent; and

30 c) comparing the Ca²⁺ concentration of step (b) with a control Ca²⁺ concentration of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC calcium channel activity.

According to another aspect of the invention, a method for identifying agents useful in the modulation of a SOC/CRAC polypeptide kinase activity is provided. The method involves:

a) contacting a SOC/CRAC polypeptide with kinase activity with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity;

b) detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and

c) comparing the kinase activity of step (b) with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC kinase activity. In some embodiments the SOC/CRAC polypeptide comprises amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24), or a fragment thereof that retains the kinase activity.

According to yet another aspect of the invention, a method for determining the level of expression of a SOC/CRAC polypeptide in a subject is provided. The method involves:

a) measuring the expression of a SOC/CRAC polypeptide in a test sample, and

b) comparing the measured expression of the SOC/CRAC polypeptide in the test sample to the expression of a SOC/CRAC polypeptide in a control containing a known level of expression to determine the level of SOC/CRAC expression in the subject. Expression is defined as SOC/CRAC mRNA expression or SOC/CRAC polypeptide expression. Various methods can be used to measure expression. The preferred embodiments of the invention utilize PCR and Northern blotting for measuring mRNA expression, and monoclonal or polyclonal SOC/CRAC antisera as reagents for measuring SOC/CRAC polypeptide expression. In preferred embodiments, the SOC/CRAC molecule (nucleic acid and/or polypeptide) is SOC-2/CRAC-1. In other preferred embodiments, the SOC/CRAC molecule is SOC-3/CRAC-2. In yet further preferred embodiments, the SOC/CRAC molecule is SOC-4/CRAC-3. In certain embodiments, the test samples include biopsy samples and biological fluids such as blood. The method is useful, e.g., for assessing the presence or absence or stage of a proliferative disorder in a subject.

The invention also contemplates kits comprising a package including assays for SOC/CRAC epitopes, SOC/CRAC nucleic acids, and instructions, and optionally related materials such as controls, for example, a number, color chart, or an epitope of the expression product of the foregoing isolated nucleic acid molecules of the invention for comparing, for

example, the level of SOC/CRAC polypeptides or SOC/CRAC nucleic acid forms (wild-type or mutant) in a test sample to the level in a control sample having a known amount of a SOC/CRAC nucleic acid or SOC/CRAC polypeptide. This comparison can be used to assess in a subject a risk of developing a cancer or the progression of a cancer. The kits may also include assays for other known genes, and expression products thereof, associated with, for example, proliferative disorders (e.g., BRCA, p53, etc.). In a preferred embodiment, the kit comprises a package containing: (a) a binding agent that selectively binds to an isolated nucleic acid of the invention or an expression product thereof to obtain a measured test value, (b) a control containing a known amount of a SOC/CRAC nucleic acid or a SOC/CRAC polypeptide to obtain a measured control value, and (c) instructions for comparing the measured test value to the measured control value to determine the amount of SOC/CRAC nucleic acid or expression product thereof in a sample.

The invention provides isolated nucleic acid molecules, unique fragments thereof, expression vectors containing the foregoing, and host cells containing the foregoing. The invention also provides isolated binding polypeptides and binding agents which bind such polypeptides, including antibodies, and pharmaceutical compositions containing any of the compositions of the invention. The foregoing can be used, *inter alia*, in the diagnosis or treatment of conditions characterized by the aberrant expression levels and/or the presence of mutant forms of a SOC/CRAC nucleic acid or polypeptide. The invention also provides methods for identifying agents that alter the function of the SOC/CRAC polypeptide.

These and other aspects of the invention, as well as various advantages and utilities, will be more apparent with reference to the detailed description of the preferred embodiments.

Brief Description of the Sequences

SEQ ID NO:1 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:2 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:1).

SEQ ID NO:3 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:4 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:3).

SEQ ID NO:5 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:6 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:5).

SEQ ID NO:7 is a partial nucleotide sequence of the mouse homologue (mSOC-2/CRAC-1) of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:8 is the predicted amino acid sequence of the translation product of the mSOC-2/CRAC-1 cDNA (SEQ ID NO:7).

5 SEQ ID NO:9 is the nucleotide sequence of the mouse MLSN-1 (SOC-1) cDNA.

SEQ ID NO:10 is the predicted amino acid sequence of the translation product of the mouse MLSN-1 (SOC-1) cDNA (SEQ ID NO:9).

SEQ ID NO:11 is the nucleotide sequence of a human calcium channel cDNA with GenBank Acc. no.: AB001535.

10 SEQ ID NO:12 is the predicted amino acid sequence of the translation product of the human calcium channel cDNA with GenBank Acc. no.: AB001535 (SEQ ID NO:11).

SEQ ID NO:13 is the amino acid sequence of a *C. Elegans* polypeptide at the c05c12.3 locus.

15 SEQ ID NO:14 is the amino acid sequence of a *C. Elegans* polypeptide at the F54D1 locus.

SEQ ID NO:15 is the amino acid sequence of a *C. Elegans* polypeptide at the t01H8 locus.

SEQ ID NO:16 is the nucleotide sequence of a mouse kidney cDNA with GenBank Acc. no.: AI226731.

20 SEQ ID NO:17 is the predicted amino acid sequence of the translation product of the mouse kidney cDNA with GenBank Acc. no.: AI226731 (SEQ ID NO:16).

SEQ ID NO:18 is the nucleotide sequence of a human brain cDNA with GenBank Acc. no.: H18835.

25 SEQ ID NO:19 is the predicted amino acid sequence of the translation product of the human brain cDNA with GenBank Acc. no.: H18835 (SEQ ID NO:18).

SEQ ID NO:20 is the nucleotide sequence of the human EST with GenBank Acc. no.: AA419592.

SEQ ID NO:21 is the nucleotide sequence of the human EST with GenBank Acc. no.: AA419407.

30 SEQ ID NO:22 is the nucleotide sequence of the mouse EST with GenBank Acc. no.: AI098310.

SEQ ID NO:23 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA that contains the SOC-2/CRAC-1 sequences of SEQ ID NO:1, SEQ ID NO:3, and SEQ ID NO:5.

SEQ ID NO:24 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:23).

SEQ ID NO:25 is a partial nucleotide sequence of the human SOC-3/CRAC-2 cDNA.

SEQ ID NO:26 is the predicted amino acid sequence of the translation product of human SOC-3/CRAC-2 cDNA (SEQ ID NO:25).

SEQ ID NO:27 is the full nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:28 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:27).

SEQ ID NO:29 is the full nucleotide sequence of the human SOC-3/CRAC-2 cDNA.

SEQ ID NO:30 is the predicted amino acid sequence of the translation product of human SOC-3/CRAC-2 cDNA (SEQ ID NO:29).

SEQ ID NO:31 is the full nucleotide sequence of the human SOC-4/CRAC-3 cDNA.

SEQ ID NO:32 is the predicted amino acid sequence of the translation product of human SOC-4/CRAC-3 cDNA (SEQ ID NO:31).

Brief Description of the Drawings

Figure 1 is a schematic depicting the intron/exon organization of the chicken SOC-2/CRAC-1 genomic sequence, as well as the putative transmembrane (TM) domains, and the targeting constructs utilized in the knockout experiments.

Detailed Description of the Invention

One aspect of the invention involves the partial cloning of cDNAs encoding members of a novel family of calcium channel polypeptides, referred to herein as "SOC/CRAC" (designated "SOC" or "CRAC" or "ICRAC", for Sore Operated Channels or Calcium Release Activated Channels, or CECH). Although not intending to be bound to any particular mechanism or theory, we believe that a SOC/CRAC family member is a transmembrane calcium channel that modulates Ca^{2+} flux "into" and "out of" a cell; in certain instances it may be activated upon depletion of Ca^{2+} from intracellular calcium stores, allowing Ca^{2+} influx into the cell.

The first three isolated SOC/CRAC members disclosed herein, define a new family of calcium channels which is distinct from previously described calcium channels, such as voltage gated calcium channels, ryanodine receptor/inositol-1,4,5-triphosphate receptor

channels, and Transient Receptor Potential (TRP) channels. The SOC/CRAC family of calcium channels exhibits high selectivity (with a P_{Ca}/P_{Na} ratio near 1000), a unitary conductance below the detection level of the patch clamp method (the conductance estimated at approximately 0.2 picosiemens), and are subject to inhibition by high intracellular calcium levels. Although not intending to be bound to any particular mechanism or theory, we believe that SOC/CRAC calcium channels are responsible for the majority of, for example, calcium entry which occurs when intracellular calcium stores are depleted, and that SOC/CRAC currents are important for initiating various types of calcium-dependent processes. Thus, we believe that SOC/CRAC calcium channels play an important role in cellular calcium homeostasis by, e.g., modulating the supply of calcium to refill intracellular stores when depleted.

The isolated full-length sequence of a representative, first member of the SOC/CRAC family, human SOC/CRAC nucleic acid (cDNA), SOC-2/CRAC-1, is represented as the nucleic acid of SEQ ID NO:27. This nucleic acid sequence codes for the SOC-2/CRAC-1 polypeptide with the predicted amino acid sequence disclosed herein as SEQ ID NO:28. A homologous mouse cDNA sequence (>90% identity to the human at the nucleotide level) is represented as the nucleic acid of SEQ ID NO:7, and codes for a unique fragment of a mouse SOC-2/CRAC-1 polypeptide having the predicted, partial amino acid sequence represented as SEQ ID NO:8. Analysis of the SOC-2/CRAC-1 partial sequence by comparison to nucleic acid and protein databases show that SOC-2/CRAC-1 shares a limited homology to mouse MLSN-1 (SOC-1, SEQ ID NOs: 9 and 10). Limited homology is also shared between SOC-2/CRAC-1 and three *C. Elegans* polypeptides (SEQ ID NOs: 13, 14, and 15). We further believe that SOC-2/CRAC-1 plays a role in the regulation of cellular Ca^{2+} fluxing and, in particular, lymphocyte Ca^{2+} fluxing.

A second member of the human SOC/CRAC family of calcium channels, SOC-3/CRAC-2, is represented as the nucleic acid of SEQ ID NO:29, and codes for the human SOC-3/CRAC-2 polypeptide having the predicted amino acid sequence represented as SEQ ID NO:30 (this molecule may also be referred to as CECH2). SOC-3/CRAC-2 is predominantly expressed in human hematopoietic cells (including peripheral blood lymphocytes, liver, bone marrow, spleen, thymus, lymph nodes, heart, and kidney. Expression can also be detected (at lesser levels) in brain, skeletal muscle colon, small intestine, placenta, lung, and cells (cell lines) such as HL-60, HeLa, K562, MOLT-4, SW-480, A459, and G361.

A third member of the human SOC/CRAC family of calcium channels, SOC-4/CRAC-3, is represented as the nucleic acid of SEQ ID NO:31, and codes for the human SOC-4/CRAC-3 polypeptide having the predicted amino acid sequence represented as SEQ ID NO:32 (this molecule may also be referred to as CECH6). It specifically expressed in the prostate gland/cells.

As used herein, a SOC/CRAC calcium channel nucleic acid (also referred to herein as a "SOC/CRAC nucleic acid" refers to a nucleic acid molecule which: (1) hybridizes under stringent conditions to one or more of the nucleic acids having the sequences of SEQ. ID NOS. 7, 27, 29, and/or 31 (sequences of the mouse and human SOC-2/CRAC-1, human SOC-3/CRAC-2, and human SOC-4/CRAC-3 nucleic acids), and (2) codes for a SOC-2/CRAC-1, a SOC-3/CRAC-2 or a SOC-4/CRAC-3 calcium channel polypeptide, respectively, or unique fragments of said SOC-2/CRAC-1, SOC-3/CRAC-2, or SOC-4/CRAC-3 polypeptide.

As used herein, a SOC/CRAC calcium channel polypeptide (also referred to herein as a "SOC/CRAC polypeptide") refers to a polypeptide that is coded for by a SOC-2/CRAC-1, a SOC-3/CRAC-2, and/or a SOC-4/CRAC-3 nucleic acid. Preferably, the above-identified SOC/CRAC polypeptides mediate transport of calcium into and out of a cell.

SOC/CRAC polypeptides also are useful as immunogenic molecules for the generation of binding polypeptides (e.g., antibodies) which bind selectively to SOC/CRAC (e.g., SOC-2/CRAC-1, SOC-3/CRAC-2, and/or SOC-4/CRAC-3) polypeptides. Such antibodies can be used in diagnostic assays to identify and/or quantify the presence of a SOC/CRAC polypeptide in a sample, such as a biological fluid or biopsy sample. SOC/CRAC polypeptides further embrace functionally equivalent fragments, variants, and analogs of the preferred SOC/CRAC polypeptides, provided that the fragments, variants, and analogs also are useful in mediating calcium transport into and out of intracellular calcium stores.

As used herein, "SOC/CRAC calcium channel activity" refers to Ca^{2+} transport ("Ca²⁺ fluxing") across the plasma membrane that is mediated by a SOC/CRAC calcium channel polypeptide. The SOC/CRAC calcium channel polypeptide typically has one or more of the following properties: high selectivity, a unitary conductance below the detection level of the patch clamp method, and are subject to inhibition by high intracellular calcium levels. Such activity can be easily detected using standard methodology well known in the art. See, e.g., the Examples and Neher, E., "Ion channels for communication between and within cells",

Science, 1992; 256:498-502; and Hoth, M., and Penner, R., "Depletion of intracellular calcium stores activates a calcium current in mast cells", Nature, 1992; 355 (6358):353-6.

According to one aspect of the invention, isolated nucleic acid molecules which code for one or more member(s) of the SOC/CRAC family of calcium channel polypeptides are provided. The isolated nucleic acid molecules are selected from the following groups:

(a) nucleic acid molecules which hybridize under stringent conditions to one or more nucleic acid molecules selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;

(b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;

(c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and

(d) complements of (a), (b) or (c).

In certain embodiments, the isolated nucleic acid molecule comprises one or more of nucleotides 1-1212 of SEQ ID NO:1; nucleotides 1-739 of SEQ ID NO:3; nucleotides 1-1579 of SEQ ID NO:5; nucleotides 1-5117 of SEQ ID NO:23; the mouse homolog for SOC-2/CRAC-1 corresponding to SEQ ID NO:7; nucleotides 1-2180 of SEQ ID NO:25; nucleotides 382-5976 of SEQ ID NO:27; nucleotides 73-3714 of SEQ ID NO:29; and nucleotides 23-3434 of SEQ ID NO:31. In yet other embodiments, the isolated nucleic acid molecule comprises a molecule which encodes a polypeptide having one or more sequences selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32.

According to yet another aspect of the invention, an isolated nucleic acid molecule is provided which is selected from the group consisting of:

(a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, (of sufficient length to represent a sequence unique within the human genome); and (b) complements of (a), provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to a sequence in the prior art as represented by the sequence group consisting of: (1) sequences having the SEQ ID NOs or GenBank accession numbers of Table I, (2) complements of (1), and (3) fragments of (1) and (2).

In some embodiments, the sequence of contiguous nucleotides is selected from the group consisting of (1) at least two contiguous nucleotides nonidentical to the sequence group, (2) at least three contiguous nucleotides nonidentical to the sequence group, (3) at least four contiguous nucleotides nonidentical to the sequence group, (4) at least five contiguous nucleotides nonidentical to the sequence group, (5) at least six contiguous nucleotides nonidentical to the sequence group, (6) at least seven contiguous nucleotides nonidentical to the sequence group.

In other embodiments, the unique fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 40 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, 200 nucleotides, 1000 nucleotides and every integer length therebetween.

According to another aspect of the invention, expression vectors and host cells containing (e.g., transformed or transfected with) expression vectors comprising the nucleic acid molecules disclosed herein operably linked to a promoter are provided. In certain preferred embodiments, the host cells are eukaryotic cells.

The isolated nucleic acid molecules disclosed herein have various utilities, including their use as probes and primers to identify additional members of the SOC/CRAC family of calcium channels, as diagnostic reagents for identifying the presence of SOC/CRAC polypeptides in biological or other samples, and as agents for generating SOC/CRAC binding polypeptides (e.g., antibodies) that can be used as reagents in diagnostic and therapeutic assays to identify the presence, absence, and/or amounts of a SOC/CRAC nucleic acid or polypeptide in a biological or other sample.

As used herein with respect to nucleic acids, the term "isolated" means: (i) amplified *in vitro* by, for example, polymerase chain reaction (PCR); (ii) recombinantly produced by cloning; (iii) purified, as by cleavage and gel separation; or (iv) synthesized by, for example, chemical synthesis. An isolated nucleic acid is one which is readily manipulatable by recombinant DNA techniques well known in the art. Thus, a nucleotide sequence contained in a vector in which 5' and 3' restriction sites are known or for which polymerase chain reaction (PCR) primer sequences have been disclosed is considered isolated but a nucleic acid sequence existing in its native state in its natural host is not. An isolated nucleic acid may be substantially purified, but need not be. For example, a nucleic acid that is isolated within a cloning or expression vector is not pure in that it may comprise only a tiny percentage of the

material in the cell in which it resides. Such a nucleic acid is isolated, however, as the term is used herein because it is readily manipulatable by standard techniques known to those of ordinary skill in the art.

As used herein with respect to polypeptides (discussed below), the term "isolated" means separated from its native environment in sufficiently pure form so that it can be manipulated or used for any one of the purposes of the invention. Thus, isolated means sufficiently pure to be used (i) to raise and/or isolate antibodies, (ii) as a reagent in an assay, or (iii) for sequencing, etc.

Homologs and alleles of the SOC/CRAC nucleic acids of the invention can be identified by conventional techniques. Thus, an aspect of the invention is those nucleic acid sequences which code for SOC/CRAC polypeptides and which hybridize to a nucleic acid molecule selected from a group consisting of the nucleic acid of SEQ ID NO:1, the nucleic acid of SEQ ID NO:3, the nucleic acid of SEQ ID NO:5, the nucleic acid of SEQ ID NO:7, the nucleic acid of SEQ ID NO:23, the nucleic acid of SEQ ID NO:25, the nucleic acid of SEQ ID NO:27, the nucleic acid of SEQ ID NO:29, and the nucleic acid of SEQ ID NO:31, under stringent conditions. The term "stringent conditions" as used herein refers to parameters with which the art is familiar. Nucleic acid hybridization parameters may be found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. More specifically, stringent conditions, as used herein, refers, for example, to hybridization at 65°C in hybridization buffer (3.5 x SSC, 0.02% Ficoll, 0.02% polyvinyl pyrrolidone, 0.02% Bovine Serum Albumin, 2.5mM NaH₂PO₄(pH7), 0.5% SDS, 2mM EDTA). SSC is 0.15M sodium chloride/0.15M sodium citrate, pH7; SDS is sodium dodecyl sulphate; and EDTA is ethylenediaminetetracetic acid. After hybridization, the membrane upon which the DNA is transferred is washed at 2 x SSC at room temperature and then at 0.1 x SSC/0.1 x SDS at temperatures up to 68°C.

There are other conditions, reagents, and so forth which can be used, and would result in a similar degree of stringency. The skilled artisan will be familiar with such conditions, and thus they are not given here. It will be understood, however, that the skilled artisan will be able to manipulate the conditions in a manner to permit the clear identification of homologs and alleles of the SOC/CRAC nucleic acids of the invention. The skilled artisan also is familiar with the methodology for screening cells and libraries for expression of such

molecules which then are routinely isolated, followed by isolation of the pertinent nucleic acid molecule and sequencing.

In general homologs and alleles typically will share at least 40% nucleotide identity and/or at least 50% amino acid identity to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and/or SEQ ID NO:31, and SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, respectively. In some instances sequences will share at least 50% nucleotide identity and/or at least 65% amino acid identity and in still other instances sequences will share at least 60% nucleotide identity and/or at least 75% amino acid identity. The homology can be calculated using various, publicly available software tools developed by NCBI (Bethesda, Maryland) that can be obtained through the internet (<ftp://ncbi.nlm.nih.gov/pub/>). Exemplary tools include the BLAST system available at <http://www.ncbi.nlm.nih.gov>. Pairwise and ClustalW alignments (BLOSUM30 matrix setting) as well as Kyte-Doolittle hydropathic analysis can be obtained using the MacVetor sequence analysis software (Oxford Molecular Group). Watson-Crick complements of the foregoing nucleic acids also are embraced by the invention.

In screening for SOC/CRAC related genes, such as homologs and alleles of SOC-2/CRAC-1 and/or SOC-3/CRAC-2, a Southern blot may be performed using the foregoing conditions, together with a radioactive probe. After washing the membrane to which the DNA is finally transferred, the membrane can be placed against X-ray film or a phosphorimager plate to detect the radioactive signal.

Given that the expression of the SOC/CRAC gene is prominent in certain human tissues (e.g., SOC-2/CRAC-1: lymphoid tissue/heart, SOC-3/CRAC-2: kidney/colon, SOC-4/CRAC-3: prostate), and given the teachings herein of partial human SOC/CRAC clones, full-length and other mammalian sequences corresponding to the human SOC/CRAC partial nucleic acid sequences can be isolated from, for example, a cDNA library prepared from one or more of the tissues in which SOC-2/CRAC-1 expression is prominent, SOC-3/CRAC-2 is prominent, and/or SOC-4/CRAC-3 expression is prominent, using standard colony hybridization techniques.

The invention also includes degenerate nucleic acids which include alternative codons to those present in the native materials. For example, serine residues are encoded by the codons TCA, AGT, TCC, TCG, TCT and AGC. Each of the six codons is equivalent for the purposes of encoding a serine residue. Thus, it will be apparent to one of ordinary skill in the

art that any of the serine-encoding nucleotide triplets may be employed to direct the protein synthesis apparatus, *in vitro* or *in vivo*, to incorporate a serine residue into an elongating SOC/CRAC polypeptide. Similarly, nucleotide sequence triplets which encode other amino acid residues include, but are not limited to: CCA, CCC, CCG and CCT (proline codons); CGA, CGC, CGG, CGT, AGA and AGG (arginine codons); ACA, ACC, ACG and ACT (threonine codons); AAC and AAT (asparagine codons); and ATA, ATC and ATT (isoleucine codons). Other amino acid residues may be encoded similarly by multiple nucleotide sequences. Thus, the invention embraces degenerate nucleic acids that differ from the biologically isolated nucleic acids in codon sequence due to the degeneracy of the genetic code.

The invention also provides isolated unique fragments of an isolated nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31. A unique fragment is one that is a 'signature' for the larger nucleic acid. For example, the unique fragment is long enough to assure that its precise sequence is not found in molecules within the human genome outside of the SOC/CRAC nucleic acids defined above (and human alleles). Those of ordinary skill in the art may apply no more than routine procedures to determine if a fragment is unique within the human genome.

Unique fragments, however, exclude fragments completely composed of the nucleotide sequences of any of GenBank accession numbers and SEQ ID NOs listed in Table I (SEQ ID NO:9, AB001535, AI226731, H18835, AA419592, AA261842, AA419407, AI098310, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853), or other previously published sequences as of the filing date of this application.

A fragment which is completely composed of the sequence described in the foregoing GenBank deposits and SEQ ID NO:9, is one which does not include any of the nucleotides unique to the sequences of the invention. Thus, a unique fragment must contain a nucleotide sequence other than the exact sequence of those in GenBank or fragments thereof. The difference may be an addition, deletion or substitution with respect to the GenBank sequence or it may be a sequence wholly separate from the GenBank sequence.

Unique fragments can be used as probes in Southern and Northern blot assays to identify such nucleic acids, or can be used in amplification assays such as those employing PCR. As known to those skilled in the art, large probes such as 200, 250, 300 or more nucleotides are preferred for certain uses such as Southern and Northern blots, while smaller fragments will be preferred for uses such as PCR. Unique fragments also can be used to produce fusion proteins for generating antibodies or determining binding of the polypeptide fragments, as demonstrated in the Examples, or for generating immunoassay components. Likewise, unique fragments can be employed to produce nonfused fragments of the SOC/CRAC polypeptides, useful, for example, in the preparation of antibodies, immunoassays or therapeutic applications. Unique fragments further can be used as antisense molecules to inhibit the expression of SOC/CRAC nucleic acids and polypeptides, respectively.

As will be recognized by those skilled in the art, the size of the unique fragment will depend upon its conservancy in the genetic code. Thus, some regions of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and complements thereof, will require longer segments to be unique while others will require only short segments, typically between 12 and 32 nucleotides long (e.g. 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32 bases) or more, up to the entire length of the disclosed sequence. As mentioned above, this disclosure intends to embrace each and every fragment of each sequence, beginning at the first nucleotide, the second nucleotide and so on, up to 8 nucleotides short of the end, and ending anywhere from nucleotide number 8, 9, 10 and so on for each sequence, up to the very last nucleotide, (provided the sequence is unique as described above). Virtually any segment of the region of SEQ ID NO:1 beginning at nucleotide 1 and ending at nucleotide 1212, or SEQ ID NO:3 beginning at nucleotide 1 and ending at nucleotide 739, or SEQ ID NO:5 beginning at nucleotide 1 and ending at nucleotide 1579, or SEQ ID NO:7 beginning at nucleotide 1 and ending at nucleotide 3532, or SEQ ID NO:23 beginning at nucleotide 1 and ending at nucleotide 5117, SEQ ID NO:25 beginning at nucleotide 1 and ending at nucleotide 2180, SEQ ID NO:27 beginning at nucleotide 1 and ending at nucleotide 7419, or SEQ ID NO:29 beginning at nucleotide 1 and ending at nucleotide 4061, or SEQ ID NO:31 beginning at nucleotide 1 and ending at nucleotide 4646, or complements thereof, that is 20 or more nucleotides in length will be unique. Those skilled in the art are well versed in methods for selecting such sequences, typically on the basis of the ability of the unique

fragment to selectively distinguish the sequence of interest from other sequences in the human genome of the fragment to those on known databases typically is all that is necessary, although *in vitro* confirmatory hybridization and sequencing analysis may be performed.

As mentioned above, the invention embraces antisense oligonucleotides that selectively bind to a nucleic acid molecule encoding a SOC/CRAC polypeptide, to decrease SOC/CRAC calcium channel activity. When using antisense preparations of the invention, slow intravenous administration is preferred.

As used herein, the term "antisense oligonucleotide" or "antisense" describes an oligonucleotide that is an oligoribonucleotide, oligodeoxyribonucleotide, modified oligoribonucleotide, or modified oligodeoxyribonucleotide which hybridizes under physiological conditions to DNA comprising a particular gene or to an mRNA transcript of that gene and, thereby, inhibits the transcription of that gene and/or the translation of that mRNA. The antisense molecules are designed so as to interfere with transcription or translation of a target gene upon hybridization with the target gene or transcript. Those skilled in the art will recognize that the exact length of the antisense oligonucleotide and its degree of complementarity with its target will depend upon the specific target selected, including the sequence of the target and the particular bases which comprise that sequence. It is preferred that the antisense oligonucleotide be constructed and arranged so as to bind selectively with the target under physiological conditions, i.e., to hybridize substantially more to the target sequence than to any other sequence in the target cell under physiological conditions. Based upon SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, or upon allelic or homologous genomic and/or cDNA sequences, one of skill in the art can easily choose and synthesize any of a number of appropriate antisense molecules for use in accordance with the present invention. In order to be sufficiently selective and potent for inhibition, such antisense oligonucleotides should comprise at least 10 and, more preferably, at least 15 consecutive bases which are complementary to the target, although in certain cases modified oligonucleotides as short as 7 bases in length have been used successfully as antisense oligonucleotides (Wagner et al., *Nat. Med.* 1(11):1116-1118, 1995). Most preferably, the antisense oligonucleotides comprise a complementary sequence of 20-30 bases. Although oligonucleotides may be chosen which are antisense to any region of the gene or mRNA transcripts, in preferred embodiments the antisense oligonucleotides correspond to N-terminal or 5' upstream sites such as translation initiation, transcription initiation or promoter sites. In

addition, 3'-untranslated regions may be targeted by antisense oligonucleotides. Targeting to mRNA splicing sites has also been used in the art but may be less preferred if alternative mRNA splicing occurs. In addition, the antisense is targeted, preferably, to sites in which mRNA secondary structure is not expected (see, e.g., Sainio et al., *Cell Mol. Neurobiol.* 14(5):439-457, 1994) and at which proteins are not expected to bind. Finally, although, SEQ ID No:1 discloses a cDNA sequence, one of ordinary skill in the art may easily derive the genomic DNA corresponding to this sequence. Thus, the present invention also provides for antisense oligonucleotides which are complementary to the genomic DNA corresponding to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31. Similarly, antisense to allelic or homologous SOC/CRAC cDNAs and genomic DNAs are enabled without undue experimentation.

In one set of embodiments, the antisense oligonucleotides of the invention may be composed of "natural" deoxyribonucleotides, ribonucleotides, or any combination thereof. That is, the 5' end of one native nucleotide and the 3' end of another native nucleotide may be covalently linked, as in natural systems, via a phosphodiester internucleoside linkage. These oligonucleotides may be prepared by art recognized methods which may be carried out manually or by an automated synthesizer. They also may be produced recombinantly by vectors.

In preferred embodiments, however, the antisense oligonucleotides of the invention also may include "modified" oligonucleotides. That is, the oligonucleotides may be modified in a number of ways which do not prevent them from hybridizing to their target but which enhance their stability or targeting or which otherwise enhance their therapeutic effectiveness.

The term "modified oligonucleotide" as used herein describes an oligonucleotide in which (1) at least two of its nucleotides are covalently linked via a synthetic internucleoside linkage (i.e., a linkage other than a phosphodiester linkage between the 5' end of one nucleotide and the 3' end of another nucleotide) and/or (2) a chemical group not normally associated with nucleic acids has been covalently attached to the oligonucleotide. Preferred synthetic internucleoside linkages are phosphorothioates, alkylphosphonates, phosphorodithioates, phosphate esters, alkylphosphonothioates, phosphoramidates, carbamates, carbonates, phosphate triesters, acetamides, carboxymethyl esters and peptides.

The term "modified oligonucleotide" also encompasses oligonucleotides with a covalently modified base and/or sugar. For example, modified oligonucleotides include

oligonucleotides having backbone sugars which are covalently attached to low molecular weight organic groups other than a hydroxyl group at the 3' position and other than a phosphate group at the 5' position. Thus modified oligonucleotides may include a 2'-O-alkylated ribose group. In addition, modified oligonucleotides may include sugars such as arabinose instead of ribose. The present invention, thus, contemplates pharmaceutical preparations containing modified antisense molecules that are complementary to and hybridizable with, under physiological conditions, nucleic acids encoding SOC/CRAC polypeptides, together with pharmaceutically acceptable carriers. Antisense oligonucleotides may be administered as part of a pharmaceutical composition. Such a pharmaceutical composition may include the antisense oligonucleotides in combination with any standard physiologically and/or pharmaceutically acceptable carriers which are known in the art. The compositions should be sterile and contain a therapeutically effective amount of the antisense oligonucleotides in a unit of weight or volume suitable for administration to a patient. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. The term "physiologically acceptable" refers to a non-toxic material that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art.

The invention also involves expression vectors coding for SOC/CRAC proteins and fragments and variants thereof and host cells containing those expression vectors. Virtually any cells, prokaryotic or eukaryotic, which can be transformed with heterologous DNA or RNA and which can be grown or maintained in culture, may be used in the practice of the invention. Examples include bacterial cells such as E.coli and eukaryotic cells such as mouse, hamster, pig, goat, primate, yeast, xenopous, etc. They may be of a wide variety of tissue types, including mast cells, fibroblasts, oocytes and lymphocytes, and they may be primary cells or cell lines. Specific examples include CHO cells and COS cells. Cell-free transcription systems also may be used in lieu of cells.

As used herein, a "vector" may be any of a number of nucleic acids into which a desired sequence may be inserted by restriction and ligation for transport between different genetic environments or for expression in a host cell. Vectors are typically composed of DNA although RNA vectors are also available. Vectors include, but are not limited to,

plasmids, phagemids and virus genomes. A cloning vector is one which is able to replicate in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence may be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence may occur many times as the plasmid increases in copy number within the host bacterium or just a single time per host before the host reproduces by mitosis. In the case of phage, replication may occur actively during a lytic phase or passively during a lysogenic phase. An expression vector is one into which a desired DNA sequence may be inserted by restriction and ligation such that it is operably joined to regulatory sequences and may be expressed as an RNA transcript. Vectors may further contain one or more marker sequences suitable for use in the identification of cells which have or have not been transformed or transfected with the vector. Markers include, for example, genes encoding proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art (e.g., β -galactosidase or alkaline phosphatase), and genes which visibly affect the phenotype of transformed or transfected cells, hosts, colonies or plaques (e.g., green fluorescent protein). Preferred vectors are those capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

As used herein, a coding sequence and regulatory sequences are said to be "operably" joined when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

The precise nature of the regulatory sequences needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 5' non-transcribed

and 5' non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. Especially, such 5' non-transcribed regulatory sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined gene.

5 Regulatory sequences may also include enhancer sequences or upstream activator sequences as desired. The vectors of the invention may optionally include 5' leader or signal sequences. The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

According to yet another aspect of the invention, isolated SOC/CRAC polypeptides

10 are provided. Preferably, the isolated SOC/CRAC polypeptides are encoded by the isolated SOC/CRAC nucleic acid molecules disclosed herein. More preferably, the isolated SOC/CRAC polypeptides of the invention are encoded by the nucleic acid molecules having SEQ ID Nos. 1, 3, 5, 7, 23, 25, 27, 29, and 31. In yet other embodiments, the isolated SOC/CRAC polypeptides of the invention have an amino acid sequence selected from the

15 group consisting of SEQ ID Nos. 2, 4, 6, 8, 24, 26, 28, 30 and 32. Preferably, the isolated SOC/CRAC polypeptides are of sufficient length to represent a sequence unique within the human genome. Thus, the preferred embodiments include a sequence of contiguous amino acids which is not identical to a prior art sequence as represented by the sequence group consisting of the contiguous amino acids identified in Table II (SEQ ID NO:10, SEQ ID

20 NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19 and GenBank Acc. Nos. AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572).

In certain embodiments, the isolated SOC/CRAC polypeptides are immunogenic and

25 can be used to generate binding polypeptides (e.g., antibodies) for use in diagnostic and therapeutic applications. Such binding polypeptides also are useful for detecting the presence, absence, and/or amounts of a SOC/CRAC nucleic acid or polypeptide in a sample such as a biological fluid or biopsy sample. Preferably, the SOC/CRAC polypeptides that are useful for generating binding polypeptides are unique polypeptides and, therefore, binding of the

30 antibody to a SOC/CRAC polypeptide in a sample is selective for the SOC/CRAC polypeptide.

Expression vectors containing all the necessary elements for expression are commercially available and known to those skilled in the art. See, e.g., Sambrook et al.,

Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989. Cells are genetically engineered by the introduction into the cells of heterologous DNA (RNA) encoding a SOC/CRAC polypeptide or fragment or variant thereof. The heterologous DNA (RNA) is placed under operable control of transcriptional elements to permit the expression of the heterologous DNA in the host cell.

Preferred systems for mRNA expression in mammalian cells are those such as pRc/CMV (available from Invitrogen, Carlsbad, CA) that contain a selectable marker such as a gene that confers G418 resistance (which facilitates the selection of stably transfected cell lines) and the human cytomegalovirus (CMV) enhancer-promoter sequences. Additionally, suitable for expression in primate or canine cell lines is the pCEP4 vector (Invitrogen, Carlsbad, CA), which contains an Epstein Barr virus (EBV) origin of replication, facilitating the maintenance of plasmid as a multicopy extrachromosomal element. Another expression vector is the pEF-BOS plasmid containing the promoter of polypeptide Elongation Factor 1 α , which stimulates efficiently transcription *in vitro*. The plasmid is described by Mishizuma and Nagata (*Nuc. Acids Res.* 18:5322, 1990), and its use in transfection experiments is disclosed by, for example, Demoulin (*Mol. Cell. Biol.* 16:4710-4716, 1996). Still another preferred expression vector is an adenovirus, described by Stratford-Perricaudet, which is defective for E1 and E3 proteins (*J. Clin. Invest.* 90:626-630, 1992). The use of the adenovirus as an Adeno.P1A recombinant is disclosed by Warnier et al., in intradermal injection in mice for immunization against P1A (*Int. J. Cancer*, 67:303-310, 1996).

The invention also embraces so-called expression kits, which allow the artisan to prepare a desired expression vector or vectors. Such expression kits include at least separate portions of each of the previously discussed coding sequences. Other components may be added, as desired, as long as the previously mentioned sequences, which are required, are included.

It will also be recognized that the invention embraces the use of the above described, SOC/CRAC cDNA sequence containing expression vectors, to transfect host cells and cell lines, by these prokaryotic (e.g., *E. coli*), or eukaryotic (e.g., CHO cells, COS cells, yeast expression systems and recombinant baculovirus expression in insect cells). Especially useful are mammalian cells such as mouse, hamster, pig, goat, primate, etc. They may be of a wide variety of tissue types, and include primary cells and cell lines. Specific examples include dendritic cells, U293 cells, peripheral blood leukocytes, bone marrow stem cells and embryonic stem cells. The invention also permits the construction of SOC/CRAC gene

"knock-outs" in cells and in animals, providing materials for studying certain aspects of SOC/CRAC calcium channel activity.

The invention also provides isolated polypeptides (including whole proteins and partial proteins), encoded by the foregoing SOC/CRAC nucleic acids, and include the polypeptides of SEQ ID NO:2, 4, 6, 8, 24, 26, 28, 30, 32, and unique fragments thereof. Such polypeptides are useful, for example, to regulate calcium transport-mediated cell growth, differentiation and proliferation, to generate antibodies, as components of immunoassays, etc. Polypeptides can be isolated from biological samples including tissue or cell homogenates, and can also be expressed recombinantly in a variety of prokaryotic and eukaryotic expression systems by constructing an expression vector appropriate to the expression system, introducing the expression vector into the expression system, and isolating the recombinantly expressed protein. Short polypeptides, including antigenic peptides (such as are presented by MHC molecules on the surface of a cell for immune recognition) also can be synthesized chemically using well-established methods of peptide synthesis.

A unique fragment of a SOC/CRAC polypeptide, in general, has the features and characteristics of unique fragments as discussed above in connection with nucleic acids. As will be recognized by those skilled in the art, the size of the unique fragment will depend upon factors such as whether the fragment constitutes a portion of a conserved protein domain. Thus, some regions of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, will require longer segments to be unique while others will require only short segments, typically between 5 and 12 amino acids (e.g. 5, 6, 7, 8, 9, 10, 11 and 12 amino acids long or more, including each integer up to the full length, >1,000 amino acids long). Virtually any segment of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, excluding the ones that share identity with it (the polypeptides identified in Table II - SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, and GenBank Acc. Nos. AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572) that is 9 or more amino acids in length will be unique.

Unique fragments of a polypeptide preferably are those fragments which retain a distinct functional capability of the polypeptide. Functional capabilities which can be retained in a unique fragment of a polypeptide include Ca^{2+} fluxing, high selectivity, a unitary

conductance below the detection level of the patch clamp method, and/or and are subject to inhibition by high intracellular calcium levels.

One important aspect of a unique fragment is its ability to act as a signature for identifying the polypeptide. Optionally, another aspect of a unique fragment is its ability to provide an immune response in an animal. Those skilled in the art are well versed in methods for selecting unique amino acid sequences, typically on the basis of the ability of the unique fragment to selectively distinguish the sequence of interest from non-family members. A comparison of the sequence of the fragment to those on known databases typically is all that is necessary.

The invention embraces variants of the SOC/CRAC polypeptides described above. As used herein, a "variant" of a SOC/CRAC polypeptide is a polypeptide which contains one or more modifications to the primary amino acid sequence of a SOC/CRAC polypeptide. Modifications which create a SOC/CRAC polypeptide variant are typically made to the nucleic acid which encodes the SOC/CRAC polypeptide, and can include deletions, point mutations, truncations, amino acid substitutions and addition of amino acids or non-amino acid moieties to: 1) reduce or eliminate a calcium channel activity of a SOC/CRAC polypeptide; 2) enhance a property of a SOC/CRAC polypeptide, such as protein stability in an expression system or the stability of protein-protein binding; 3) provide a novel activity or property to a SOC/CRAC polypeptide, such as addition of an antigenic epitope or addition of a detectable moiety; or 4) to provide equivalent or better binding to a SOC/CRAC polypeptide receptor or other molecule. Alternatively, modifications can be made directly to the polypeptide, such as by cleavage, addition of a linker molecule, addition of a detectable moiety, such as biotin, addition of a fatty acid, and the like. Modifications also embrace fusion proteins comprising all or part of the SOC/CRAC amino acid sequence. One of skill in the art will be familiar with methods for predicting the effect on protein conformation of a change in protein sequence, and can thus "design" a variant SOC/CRAC polypeptide according to known methods. One example of such a method is described by Dahiyat and Mayo in *Science* 278:82-87, 1997, whereby proteins can be designed *de novo*. The method can be applied to a known protein to vary only a portion of the polypeptide sequence. By applying the computational methods of Dahiyat and Mayo, specific variants of a SOC/CRAC calcium channel polypeptide can be proposed and tested to determine whether the variant retains a desired conformation.

5 Variants can include SOC/CRAC polypeptides which are modified specifically to alter a feature of the polypeptide unrelated to its physiological activity. For example, cysteine residues can be substituted or deleted to prevent unwanted disulfide linkages. Similarly, certain amino acids can be changed to enhance expression of a SOC/CRAC polypeptide by eliminating proteolysis by proteases in an expression system (e.g., dibasic amino acid residues in yeast expression systems in which KEX2 protease activity is present).

10 Mutations of a nucleic acid which encodes a SOC/CRAC polypeptide preferably preserve the amino acid reading frame of the coding sequence and, preferably, do not create regions in the nucleic acid which are likely to hybridize to form secondary structures, such as hairpins or loops, which can be deleterious to expression of the variant polypeptide.

15 Mutations can be made by selecting an amino acid substitution, or by random mutagenesis of a selected site in a nucleic acid which encodes the polypeptide. Variant polypeptides are then expressed and tested for one or more activities to determine which mutation provides a variant polypeptide with the desired properties. Further mutations can be made to variants (or to non-variant SOC/CRAC polypeptides) which are silent as to the amino acid sequence of the polypeptide, but which provide preferred codons for translation in a particular host. The preferred codons for translation of a nucleic acid in, e.g., *E. coli*, are well known to those of ordinary skill in the art. Still other mutations can be made to the noncoding sequences of a SOC/CRAC gene or cDNA clone to enhance expression of the polypeptide.

20 The skilled artisan will realize that conservative amino acid substitutions may be made in SOC/CRAC polypeptides to provide functionally equivalent variants of the foregoing polypeptides, i.e., the variants retain the functional capabilities of the SOC/CRAC polypeptides. As used herein, a "conservative amino acid substitution" refers to an amino acid substitution which does not alter the relative charge or size characteristics of the protein in which the amino acid substitution is made. Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Exemplary functionally equivalent variants of the SOC/CRAC polypeptides include conservative amino acid substitutions of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32. Conservative substitutions of amino acids

include substitutions made amongst amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D.

Thus functionally equivalent variants of SOC/CRAC polypeptides, i.e., variants of SOC/CRAC polypeptides which retain the function of the natural SOC/CRAC polypeptides, are contemplated by the invention. Conservative amino-acid substitutions in the amino acid sequence of SOC/CRAC polypeptides to produce functionally equivalent variants of SOC/CRAC polypeptides typically are made by alteration of a nucleic acid encoding SOC/CRAC polypeptides (e.g., SEQ ID NOs:1, 3, 5, 7, 23, 25, 27, 29, 31). Such substitutions can be made by a variety of methods known to one of ordinary skill in the art. For example, amino acid substitutions may be made by PCR-directed mutation, site-directed mutagenesis according to the method of Kunkel (Kunkel, *Proc. Nat. Acad. Sci. U.S.A.* 82: 488-492, 1985), or by chemical synthesis of a gene encoding a SOC/CRAC polypeptide. The activity of functionally equivalent fragments of SOC/CRAC polypeptides can be tested by cloning the gene encoding the altered SOC/CRAC polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the altered SOC/CRAC polypeptide, and testing for a functional capability of the SOC/CRAC polypeptides as disclosed herein (e.g., SOC/CRAC calcium channel activity).

The invention as described herein has a number of uses, some of which are described elsewhere herein. First, the invention permits isolation of SOC/CRAC polypeptides, including the isolation of the complete SOC/CRAC polypeptide. A variety of methodologies well-known to the skilled practitioner can be utilized to obtain isolated SOC/CRAC molecules. The polypeptide may be purified from cells which naturally produce the polypeptide by chromatographic means or immunological recognition. Alternatively, an expression vector may be introduced into cells to cause production of the polypeptide. In another method, mRNA transcripts may be microinjected or otherwise introduced into cells to cause production of the encoded polypeptide. Translation of SOC/CRAC mRNA in cell-free extracts such as the reticulocyte lysate system also may be used to produce SOC/CRAC polypeptides. Those skilled in the art also can readily follow known methods for isolating SOC/CRAC polypeptides. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography and immune-affinity chromatography.

The invention also provides, in certain embodiments, "dominant negative" polypeptides derived from SOC/CRAC polypeptides. A dominant negative polypeptide is an

inactive variant of a protein, which, by interacting with the cellular machinery, displaces an active protein from its interaction with the cellular machinery or competes with the active protein, thereby reducing the effect of the active protein. For example, a dominant negative receptor which binds a ligand but does not transmit a signal in response to binding of the ligand can reduce the biological effect of expression of the ligand. Likewise, a dominant negative inactive SOC/CRAC calcium channel which interacts normally with the cell membrane but which does not mediate calcium transport can reduce calcium transport in a cell. Similarly, a dominant negative transcription factor which binds to a promoter site in the control region of a gene but does not increase gene transcription can reduce the effect of a normal transcription factor by occupying promoter binding sites without increasing transcription.

The end result of the expression of a dominant negative polypeptide in a cell is a reduction in function of active proteins. One of ordinary skill in the art can assess the potential for a dominant negative variant of a protein, and using standard mutagenesis techniques to create one or more dominant negative variant polypeptides. See, e.g., U.S. Patent No. 5,580,723 and Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press. The skilled artisan then can test the population of mutagenized polypeptides for diminution in a selected and/or for retention of such an activity. Other similar methods for creating and testing dominant negative variants of a protein will be apparent to one of ordinary skill in the art.

According to another aspect, the invention provides a method for isolating a SOC/CRAC molecule having SOC/CRAC calcium channel activity. The method involves contacting a binding molecule that is a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing one or more SOC/CRAC molecules under conditions that allow such binding (see earlier discussion) to form a complex, detecting the presence of the complex, isolating the SOC/CRAC molecule from the complex, and determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity. Thus, the invention is useful for identifying and isolating full length complementary (cDNA) or genomic nucleic acids encoding SOC/CRAC polypeptides having SOC/CRAC calcium channel activity. Identification and isolation of such nucleic acids and polypeptides may be accomplished by hybridizing/binding, under appropriate conditions well known in the art, libraries and/or restriction enzyme-digested human nucleic acids, with a labeled SOC/CRAC molecular probe. As used herein, a "label" includes molecules that are incorporated into, for

example, a SOC/CRAC molecule (nucleic acid or peptide), that can be directly or indirectly detected. A wide variety of detectable labels are well known in the art that can be used, and include labels that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc), or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseshoe peroxidase, etc.). The label may be bound to a SOC/CRAC binding partner, or incorporated into the structure of the binding partner.

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly detected through optical or electron density, radioactive emissions, nonradioactive energy transfers, etc. or indirectly detected with antibody conjugates, streptavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art. Once a library clone or hybridizing fragment is identified in the hybridization/binding reaction, it can be further isolated by employing standard isolation/cloning techniques known to those of skill in the art. See, generally, Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press. In addition, nucleic acid amplification techniques well known in the art, may also be used to locate splice variants of calcium channel (or calcium channel subunits) with SOC/CRAC calcium channel activity. Size and sequence determinations of the amplification products can reveal splice variants.

The foregoing isolated nucleic acids and polypeptides may then be compared to the nucleic acids and polypeptides of the present invention in order to identify homogeneity or divergence of the sequences, and be further characterized functionally to determine whether they belong to a family of molecules with SOC/CRAC calcium channel activity (for methodology see under the Examples section).

The isolation of the SOC/CRAC cDNA and/or partial sequences thereof also makes it possible for the artisan to diagnose a disorder characterized by an aberrant expression of SOC/CRAC. These methods involve determining expression of the SOC/CRAC gene, and/or SOC/CRAC polypeptides derived therefrom. In the former situation, such determinations can be carried out via any standard nucleic acid determination assay, including the polymerase chain reaction, or assaying with labeled hybridization probes as exemplified below. In the latter situation, such determination can be carried out via any standard immunological assay using, for example, antibodies which bind to the SOC/CRAC protein.

The invention also embraces isolated peptide binding agents which, for example, can be antibodies or fragments of antibodies ("binding polypeptides"), having the ability to selectively bind to SOC/CRAC polypeptides. Antibodies include polyclonal and monoclonal antibodies, prepared according to conventional methodology. In certain embodiments, the invention excludes binding agents (e.g., antibodies) that bind to the polypeptides encoded by the nucleic acids of SEQ ID NOs: 10, 12, 13, 14, 15, 17, and 19.

Significantly, as is well-known in the art, only a small portion of an antibody molecule, the paratope, is involved in the binding of the antibody to its epitope (see, in general, Clark, W.R. (1986) The Experimental Foundations of Modern Immunology Wiley & Sons, Inc., New York; Roitt, I. (1991) Essential Immunology, 7th Ed., Blackwell Scientific Publications, Oxford). The pFc' and Fc regions, for example, are effectors of the complement cascade but are not involved in antigen binding. An antibody from which the pFc' region has been enzymatically cleaved, or which has been produced without the pFc' region, designated an F(ab')₂ fragment, retains both of the antigen binding sites of an intact antibody. Similarly, an antibody from which the Fc region has been enzymatically cleaved, or which has been produced without the Fc region, designated an Fab fragment, retains one of the antigen binding sites of an intact antibody molecule. Proceeding further, Fab fragments consist of a covalently bound antibody light chain and a portion of the antibody heavy chain denoted Fd. The Fd fragments are the major determinant of antibody specificity (a single Fd fragment may be associated with up to ten different light chains without altering antibody specificity) and Fd fragments retain epitope-binding ability in isolation.

Within the antigen-binding portion of an antibody, as is well-known in the art, there are complementarity determining regions (CDRs), which directly interact with the epitope of the antigen, and framework regions (FRs), which maintain the tertiary structure of the paratope (see, in general, Clark, 1986; Roitt, 1991). In both the heavy chain Fd fragment and the light chain of IgG immunoglobulins, there are four framework regions (FR1 through FR4) separated respectively by three complementarity determining regions (CDR1 through CDR3). The CDRs, and in particular the CDR3 regions, and more particularly the heavy chain CDR3, are largely responsible for antibody specificity.

It is now well-established in the art that the non-CDR regions of a mammalian antibody may be replaced with similar regions of conspecific or heterospecific antibodies while retaining the epitopic specificity of the original antibody. This is most clearly manifested in the development and use of "humanized" antibodies in which non-human CDRs

are covalently joined to human FR and/or Fc/pFc' regions to produce a functional antibody. Thus, for example, PCT International Publication Number WO 92/04381 teaches the production and use of humanized murine RSV antibodies in which at least a portion of the murine FR regions have been replaced by FR regions of human origin. Such antibodies, including fragments of intact antibodies with antigen-binding ability, are often referred to as "chimeric" antibodies.

Thus, as will be apparent to one of ordinary skill in the art, the present invention also provides for F(ab')₂, Fab, Fv and Fd fragments; chimeric antibodies in which the Fc and/or FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric F(ab')₂ fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric Fab fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; and chimeric Fd fragment antibodies in which the FR and/or CDR1 and/or CDR2 regions have been replaced by homologous human or non-human sequences. The present invention also includes so-called single chain antibodies.

Thus, the invention involves binding polypeptides of numerous size and type that bind selectively to SOC/CRAC polypeptides, and complexes containing SOC/CRAC polypeptides. These binding polypeptides also may be derived also from sources other than antibody technology. For example, such polypeptide binding agents can be provided by degenerate peptide libraries which can be readily prepared in solution, in immobilized form, as bacterial flagella peptide display libraries or as phage display libraries. Combinatorial libraries also can be synthesized of peptides containing one or more amino acids. Libraries further can be synthesized of peptides and non-peptide synthetic moieties.

Phage display can be particularly effective in identifying binding peptides useful according to the invention. Briefly, one prepares a phage library (using e.g. m13, fd, or lambda phage), displaying inserts from 4 to about 80 amino acid residues using conventional procedures. The inserts may represent, for example, a completely degenerate or biased array. One then can select phage-bearing inserts which bind to the SOC/CRAC polypeptide or a complex containing a SOC/CRAC polypeptide. This process can be repeated through several cycles of reselection of phage that bind to the SOC/CRAC polypeptide or complex. Repeated rounds lead to enrichment of phage bearing particular sequences. DNA sequence analysis can be conducted to identify the sequences of the expressed polypeptides. The minimal linear

portion of the sequence that binds to the SOC/CRAC polypeptide or complex can be determined. One can repeat the procedure using a biased library containing inserts containing part or all of the minimal linear portion plus one or more additional degenerate residues upstream or downstream thereof. Yeast two-hybrid screening methods also may be used to
5 identify polypeptides that bind to the SOC/CRAC polypeptides. Thus, the SOC/CRAC polypeptides of the invention, or a fragment thereof, or complexes of SOC/CRAC can be used to screen peptide libraries, including phage display libraries, to identify and select peptide binding polypeptides that selectively bind to the SOC/CRAC polypeptides of the invention. Such molecules can be used, as described, for screening assays, for purification protocols, for
10 interfering directly with the functioning of SOC/CRAC and for other purposes that will be apparent to those of ordinary skill in the art.

A SOC/CRAC polypeptide, or a fragment thereof, also can be used to isolate naturally occurring, polypeptide binding partners which may associate with the SOC/CRAC polypeptide in the membrane of a cell. Isolation of binding partners may be performed
15 according to well-known methods. For example, isolated SOC/CRAC polypeptides can be attached to a substrate, and then a solution suspected of containing an SOC/CRAC binding partner may be applied to the substrate. If the binding partner for SOC/CRAC polypeptides is present in the solution, then it will bind to the substrate-bound SOC/CRAC polypeptide. The binding partner then may be isolated. Other proteins which are binding partners for
20 SOC/CRAC, may be isolated by similar methods without undue experimentation.

The invention also provides novel kits which could be used to measure the levels of the nucleic acids of the invention, expression products of the invention or anti-SOC/CRAC antibodies. In the case of nucleic acid detection, pairs of primers for amplifying SOC/CRAC nucleic acids can be included. The preferred kits would include controls such as known
25 amounts of nucleic acid probes, SOC/CRAC epitopes (such as SOC/CRAC expression products) or anti-SOC/CRAC antibodies, as well as instructions or other printed material. In certain embodiments the printed material can characterize risk of developing a disorder that is characterized by aberrant SOC/CRAC polypeptide expression based upon the outcome of the assay. The reagents may be packaged in containers and/or coated on wells in predetermined
30 amounts, and the kits may include standard materials such as labeled immunological reagents (such as labeled anti-IgG antibodies) and the like. One kit is a packaged polystyrene microtiter plate coated with a SOC/CRAC polypeptide and a container containing labeled anti-human IgG antibodies. A well of the plate is contacted with, for example, serum, washed

and then contacted with the anti-IgG antibody. The label is then detected. A kit embodying features of the present invention is comprised of the following major elements: packaging an agent of the invention, a control agent, and instructions. Packaging is a box-like structure for holding a vial (or number of vials) containing an agent of the invention. a vial (or number of vials) containing a control agent, and instructions. Individuals skilled in the art can readily modify packaging to suit individual needs.

Another aspect of the invention is a method for determining the level of SOC/CRAC expression in a subject. As used herein, a subject is a human, non-human primate, cow, horse, pig, sheep, goat, dog, cat or rodent. In all embodiments, human subjects are preferred. Expression is defined either as SOC/CRAC mRNA expression or SOC/CRAC polypeptide expression. Various methods can be used to measure expression. Preferred embodiments of the invention include PCR and Northern blotting for measuring mRNA expression, and monoclonal or polyclonal SOC/CRAC antisera as reagents to measure SOC/CRAC polypeptide expression. In certain embodiments, test samples such as biopsy samples, and biological fluids such as blood, are used as test samples. SOC/CRAC expression in a test sample of a subject is compared to SOC/CRAC expression in control sample to, e.g., assess the presence or absence or stage of a proliferative disorder (e.g., a lymphocyte proliferative disorder) in a subject.

SOC/CRAC polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts. Recombinantly produced SOC/CRAC polypeptides include chimeric proteins comprising a fusion of a SOC/CRAC protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the SOC/CRAC polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein. A polypeptide fused to a SOC/CRAC polypeptide or fragment may also provide means of readily detecting the fusion protein, e.g., by immunological recognition or by fluorescent labeling.

The invention is also useful in the generation of transgenic non-human animals. As used herein, "transgenic non-human animals" includes non-human animals having one or more exogenous nucleic acid molecules incorporated in germ line cells and/or somatic cells. Thus the transgenic animal include "knockout" animals having a homozygous or heterozygous gene disruption by homologous recombination, animals having episomal or chromosomally incorporated expression vectors, etc. Knockout animals can be prepared by

homologous recombination using embryonic stem cells as is well known in the art. The recombination may be facilitated using, for example, the cre/lox system or other recombinase systems known to one of ordinary skill in the art. In certain embodiments, the recombinase system itself is expressed conditionally, for example, in certain tissues or cell types, at certain embryonic or post-embryonic developmental stages, inducibly by the addition of a compound which increases or decreases expression, and the like. In general, the conditional expression vectors used in such systems use a variety of promoters which confer the desired gene expression pattern (e.g., temporal or spatial). Conditional promoters also can be operably linked to SOC/CRAC nucleic acid molecules to increase expression of SOC/CRAC in a regulated or conditional manner. *Trans*-acting negative regulators of SOC/CRAC calcium channel activity or expression also can be operably linked to a conditional promoter as described above. Such *trans*-acting regulators include antisense SOC/CRAC nucleic acids molecules, nucleic acid molecules which encode dominant negative SOC/CRAC molecules, ribozyme molecules specific for SOC/CRAC nucleic acids, and the like. The transgenic non-human animals are useful in experiments directed toward testing biochemical or physiological effects of diagnostics or therapeutics for conditions characterized by increased or decreased SOC/CRAC expression. Other uses will be apparent to one of ordinary skill in the art.

The invention further provides efficient methods of identifying agents or lead compounds for agents active at the level of a SOC/CRAC polypeptide (e.g., a SOC/CRAC polypeptide) or SOC/CRAC fragment dependent cellular function. In particular, such functions include interaction with other polypeptides or fragments thereof, and selective binding to certain molecules (e.g., agonists and antagonists). Generally, the screening methods involve assaying for compounds which interfere with SOC/CRAC calcium channel activity, although compounds which enhance SOC/CRAC calcium channel activity also can be assayed using the screening methods. Such methods are adaptable to automated, high throughput screening of compounds. The target therapeutic indications for pharmacological agents detected by the screening methods are limited only in that the target cellular function be subject to modulation by alteration of the formation of a complex comprising a SOC/CRAC polypeptide or fragment thereof and one or more SOC/CRAC binding targets. Target indications include cellular processes modulated by SOC/CRAC such as Ca^{2+} fluxing, and affected by SOC/CRAC ability to form complexes with other molecules and polypeptides as, for example, may be present in the cell membrane.

A wide variety of assays for pharmacological agents are provided, including, expression assays, labeled *in vitro* protein-protein binding assays, electrophoretic mobility shift assays, immunoassays, cell-based assays such as calcium transport assays, etc. For example, two-hybrid screens are used to rapidly examine the effect of transfected nucleic acids on the intracellular binding of SOC/CRAC or SOC/CRAC fragments to specific intracellular targets (e.g. a tyrosine kinase). The transfected nucleic acids can encode, for example, combinatorial peptide libraries or cDNA libraries. Convenient reagents for such assays, e.g., GAL4 fusion proteins, are known in the art. An exemplary cell-based assay involves transfecting a cell with a nucleic acid encoding a SOC/CRAC polypeptide fused to a GAL4 DNA binding domain and a nucleic acid encoding a reporter gene operably linked to a gene expression regulatory region, such as one or more GAL4 binding sites. Activation of reporter gene transcription occurs when the SOC/CRAC and reporter fusion polypeptides bind such as to enable transcription of the reporter gene. Agents which modulate a SOC/CRAC polypeptide mediated cell function are then detected through a change in the expression of reporter gene. Methods for determining changes in the expression of a reporter gene are known in the art.

In an expression system, for example, a SOC/CRAC polypeptide is attached to a membrane, the membrane preferably separating two fluid environments and being otherwise not permeable to Ca^{2+} . Such separation is preferred so that a change in Ca^{2+} concentration on either side of the membrane is mediated only through the attached SOC/CRAC polypeptide. Preferably, a SOC/CRAC polypeptide is expressed in an intact cell and is present on the cell-membrane (as in physiologic conditions). The cell expressing the SOC/CRAC polypeptide is preferably a eukaryotic cell, and the SOC/CRAC polypeptide is preferably recombinantly expressed, although cells naturally expressing a SOC/CRAC polypeptide may also be used. Synthetic membranes, however, containing SOC/CRAC polypeptides may also be used. See, e.g., K. Kiselyov, et al., Functional interaction between InsP3 receptors and store-operated Htrp3 channels, Nature 396, 478-82 (1998).

The cell expressing the SOC/CRAC polypeptide is incubated under conditions which, in the absence of the candidate agent, permit calcium flux into the cell and allow detection of a reference calcium concentration. For example, depletion of intracellular calcium stores with thapsigargin or other agents (Putney, J.W. Jr., in Capacitative Calcium Entry, R.G. Landes Co. and Chapman & Hall, 1997) would produce a given level of SOC/CRAC channel activation and a given reference calcium concentration. Detection of a decrease in the

foregoing activities (i.e., a decrease in the intracellular calcium concentration) relative to the reference calcium concentration indicates that the candidate agent is a lead compound for an agent to inhibit SOC/CRAC calcium channel activity. Preferred SOC/CRAC polypeptides include the polypeptides of claim 15.

5 SOC/CRAC fragments used in the methods, when not produced by a transfected nucleic acid are added to an assay mixture as an isolated polypeptide. SOC/CRAC polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts or chemically synthesized. Recombinantly produced SOC/CRAC polypeptides include chimeric proteins comprising a fusion of a SOC/CRAC
10 protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the SOC/CRAC polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein or Flag epitope.

The assay mixture is comprised of a SOC/CRAC polypeptide binding target
15 (candidate agent) capable of interacting with a SOC/CRAC polypeptide. While natural SOC/CRAC binding targets may be used, it is frequently preferred to use portions (e.g., peptides or nucleic acid fragments) or analogs (i.e., agents which mimic the SOC/CRAC binding properties of the natural binding target for purposes of the assay) of the SOC/CRAC binding target so long as the portion or analog provides binding affinity and avidity to the
20 SOC/CRAC polypeptide (or fragment thereof) measurable in the assay.

The assay mixture also comprises a candidate agent (binding target, e.g., agonist/antagonist). Typically, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a different response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration of agent or
25 at a concentration of agent below the limits of assay detection. Candidate agents encompass numerous chemical classes, although typically they are organic compounds. Preferably, the candidate agents are small organic compounds, i.e., those having a molecular weight of more than 50 yet less than about 2500, preferably less than about 1000 and, more preferably, less than about 500. Candidate agents comprise functional chemical groups necessary for
30 structural interactions with polypeptides and/or nucleic acids, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups and more preferably at least three of the functional chemical groups. The candidate agents can comprise cyclic carbon or heterocyclic structure and/or aromatic or

polyaromatic structures substituted with one or more of the above-identified functional groups. Candidate agents also can be biomolecules such as peptides, saccharides, fatty acids, sterols, isoprenoids, purines, pyrimidines, derivatives or structural analogs of the above, or combinations thereof and the like. Where the agent is a nucleic acid, the agent typically is a DNA or RNA molecule, although modified nucleic acids as defined herein are also contemplated.

Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides, synthetic organic combinatorial libraries, phage display libraries of random peptides, and the like. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural and synthetically produced libraries and compounds can be readily modified through conventional chemical, physical, and biochemical means. Further, known agents may be subjected to directed or random chemical modifications such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs of the agents. Non-SOC/CRAC calcium channel agonists and antagonists, for example, include agents such as dihydropyridines (DHPs), phenylalkylamines, omega conotoxin (omega-CgTx) and pyrazonoylguanidines.

A variety of other reagents also can be included in the mixture. These include reagents such as salts, buffers, neutral proteins (e.g., albumin), detergents, etc. which may be used to facilitate optimal protein-protein, protein-nucleic acid, and/or protein/membrane component binding association. Such a reagent may also reduce non-specific or background interactions of the reaction components. Other reagents that improve the efficiency of the assay such as protease, inhibitors, nuclease inhibitors, antimicrobial agents, and the like may also be used.

The mixture of the foregoing assay materials is incubated under conditions whereby, but for the presence of the candidate agent, the SOC/CRAC polypeptide specifically binds the cellular binding target, a portion thereof or analog thereof. The order of addition of components, incubation temperature, time of incubation, and other perimeters of the assay may be readily determined. Such experimentation merely involves optimization of the assay parameters, not the fundamental composition of the assay. Incubation temperatures typically

are between 4°C and 40°C. Incubation times preferably are minimized to facilitate rapid, high throughput screening, and typically are between 0.1 and 10 hours.

After incubation, the presence or absence of specific binding between the SOC/CRAC polypeptide and one or more binding targets is detected by any convenient method available to the user. For cell free binding type assays, a separation step is often used to separate bound from unbound components. The separation step may be accomplished in a variety of ways. Conveniently, at least one of the components is immobilized on a solid substrate, from which the unbound components may be easily separated. The solid substrate can be made of a wide variety of materials and in a wide variety of shapes, e.g., microtiter plate, microbead, dipstick, resin particle, etc. The substrate preferably is chosen to maximum signal to noise ratios, primarily to minimize background binding, as well as for ease of separation and cost.

Separation may be effected for example, by removing a bead or dipstick from a reservoir, emptying or diluting a reservoir such as a microtiter plate well, rinsing a bead, particle, chromatographic column or filter with a wash solution or solvent. The separation step preferably includes multiple rinses or washes. For example, when the solid substrate is a microtiter plate, the wells may be washed several times with a washing solution, which typically includes those components of the incubation mixture that do not participate in specific bindings such as salts, buffer, detergent, non-specific protein, etc. Where the solid substrate is a magnetic bead, the beads may be washed one or more times with a washing solution and isolated using a magnet.

Detection may be effected in any convenient way for cell-based assays such as two- or three-hybrid screens. The transcript resulting from a reporter gene transcription assay of SOC/CRAC polypeptide interacting with a target molecule typically encodes a directly or indirectly detectable product, e.g., β -galactosidase activity, luciferase activity, and the like. For cell-free binding assays, one of the components usually comprises, or is coupled to, a detectable label. A wide variety of labels can be used, such as those that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc.) or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseradish peroxidase, etc.). The label may be bound to a SOC/CRAC binding partner, or incorporated into the structure of the binding partner.

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly

detected through optical or electron density, radioactive emissions, nonradiative energy transfers, etc. or indirectly detected with antibody conjugates, strepavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art.

Of particular importance in any of the foregoing assays and binding studies is the use of a specific sequence motif identified in the SOC-2/CRAC-1 polypeptide sequence as a kinase catalytic domain. According to the invention, amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24) (or a fragment thereof), show a localized homology with the catalytic domains of eukaryotic elongation factor-2 kinase (eEF-2 kinase, GenBank Acc. no. U93850) and *Dictyostelium* myocin heavy chain kinase A (MHCK A, GenBank Acc. no. U16856), as disclosed in Ryazanov AG, et al., *Proc Natl Acad Sci U S A*, 1997, 94(10):4884-4889. Therefore, according to the invention, a method for identifying agents useful in the modulation of SOC/CRAC polypeptide kinase activity is provided. The method involves contacting a SOC/CRAC polypeptide with kinase activity, that includes, for example, amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24) with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity; detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and comparing the kinase activity in the previous step with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC kinase activity. Other controls for kinase activity can also be performed at the same time, for example, by utilizing eEF-2 kinase and/or *Dictyostelium* MHC Kinase A, in a similar manner to the SOC/CRAC member. Methods for performing such kinase activity assays are well known in the art.

The invention thus provides SOC/CRAC-specific binding agents, methods of identifying and making such agents, and their use in diagnosis, therapy and pharmaceutical development. For example, SOC/CRAC-specific agents are useful in a variety of diagnostic and therapeutic applications, especially where disease or disease prognosis is associated with altered SOC/CRAC and SOC/CRAC calcium channel fluxing characteristics. Novel SOC/CRAC-specific binding agents include SOC/CRAC-specific antibodies and other natural intracellular and extracellular binding agents identified with assays such as two hybrid screens, and non-natural intracellular and extracellular binding agents identified in screens of chemical libraries and the like.

In general, the specificity of SOC/CRAC binding to a specific molecule is determined by binding equilibrium constants. Targets which are capable of selectively binding a SOC/CRAC polypeptide preferably have binding equilibrium constants of at least about 10^7 M^{-1} , more preferably at least about 10^8 M^{-1} , and most preferably at least about 10^9 M^{-1} . The wide variety of cell based and cell free assays may be used to demonstrate SOC/CRAC-specific binding. Cell based assays include one, two and three hybrid screens, assays in which SOC/CRAC-mediated transcription is inhibited or increased, etc. Cell free assays include SOC/CRAC-protein binding assays, immunoassays, etc. Other assays useful for screening agents which bind SOC/CRAC polypeptides include fluorescence resonance energy transfer (FRET), and electrophoretic mobility shift analysis (EMSA).

Various techniques may be employed for introducing nucleic acids of the invention into cells, depending on whether the nucleic acids are introduced *in vitro* or *in vivo* in a host. Such techniques include transfection of nucleic acid- $CaPO_4$ precipitates, transfection of nucleic acids associated with DEAE, transfection with a retrovirus including the nucleic acid of interest, liposome mediated transfection, and the like. For certain uses, it is preferred to target the nucleic acid to particular cells. In such instances, a vehicle used for delivering a nucleic acid of the invention into a cell (e.g., a retrovirus, or other virus; a liposome) can have a targeting molecule attached thereto. For example, a molecule such as an antibody specific for a surface membrane protein on the target cell or a ligand for a receptor on the target cell can be bound to or incorporated within the nucleic acid delivery vehicle. For example, where liposomes are employed to deliver the nucleic acids of the invention, proteins which bind to a surface membrane protein associated with endocytosis may be incorporated into the liposome formulation for targeting and/or to facilitate uptake. Such proteins include capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half life, and the like. Polymeric delivery systems also have been used successfully to deliver nucleic acids into cells, as is known by those skilled in the art. Such systems even permit oral delivery of nucleic acids.

Other delivery systems can include time-release, delayed release or sustained release delivery systems. Such systems can avoid repeated administrations of the anti-inflammatory agent, increasing convenience to the subject and the physician. Many types of release delivery systems are available and known to those of ordinary skill in the art. They include polymer base systems such as poly(lactide-glycolide), copolyoxalates, polycaprolactones,

polyesteramides, polyorthoesters, polyhydroxybutyric acid, and polyanhydrides. Microcapsules of the foregoing polymers containing drugs are described in, for example, U.S. Patent 5,075,109. Delivery systems also include non-polymer systems that are: lipids including sterols such as cholesterol, cholesterol esters and fatty acids or neutral fats such as mono- di- and tri-glycerides; hydrogel release systems; sylastic systems; peptide based systems; wax coatings; compressed tablets using conventional binders and excipients; partially fused implants; and the like. Specific examples include, but are not limited to: (a) erosional systems in which an agent of the invention is contained in a form within a matrix such as those described in U.S. Patent Nos. 4,452,775, 4,675,189, and 5,736,152, and (b) diffusional systems in which an active component permeates at a controlled rate from a polymer such as described in U.S. Patent Nos. 3,854,480, 5,133,974 and 5,407,686. In addition, pump-based hardware delivery systems can be used, some of which are adapted for implantation.

Use of a long-term sustained release implant may be particularly suitable for treatment of chronic conditions. Long-term release, as used herein, means that the implant is constructed and arranged to delivery therapeutic levels of the active ingredient for at least 30 days, and preferably 60 days. Long-term sustained release implants are well-known to those of ordinary skill in the art and include some of the release systems described above.

The invention also contemplates gene therapy. The procedure for performing *ex vivo* gene therapy is outlined in U.S. Patent 5,399,346 and in exhibits submitted in the file history of that patent, all of which are publicly available documents. In general, it involves introduction *in vitro* of a functional copy of a gene into a cell(s) of a subject which contains a defective copy of the gene, and returning the genetically engineered cell(s) to the subject. The functional copy of the gene is under operable control of regulatory elements which permit expression of the gene in the genetically engineered cell(s). Numerous transfection and transduction techniques as well as appropriate expression vectors are well known to those of ordinary skill in the art, some of which are described in PCT application WO95/00654. *In vivo* gene therapy using vectors such as adenovirus, retroviruses, herpes virus, and targeted liposomes also is contemplated according to the invention. See, e.g., U.S. Patent Nos. 5,670,488, entitled "Adenovirus Vector for Gene Therapy", issued to Gregory et al., and 5,672,344, entitled "Viral-Mediated Gene Transfer System", issued to Kelley et al.

The invention will be more fully understood by reference to the following examples. These examples, however, are merely intended to illustrate the embodiments of the invention and are not to be construed to limit the scope of the invention.

Examples

5 As an initial approach to identifying SOC/CRAC channels, we considered publicly available data and hypothesized that the following characteristics are likely to be exhibited by SOC/CRAC calcium channels: i) SOC/CRAC calcium channels would be integral membrane proteins related (probably distantly) to one of the known calcium channel families (e.g. voltage gated, ligand gated, Trp), and therefore should have a pore region formed by a
10 tetramer of 6-7 transmembrane (TM) regions; ii) high calcium selectivity was likely to come at the price of complexity, and therefore these were likely to be large proteins; iii) the high calcium selectivity of this type of channel was likely to be useful and, therefore, highly conserved; and iv) these channels should be expressed in one or more types of lymphocytes, since ICRAC is best defined in those cell types. Since the full genome of the nematode *C. elegans* is nearing completion, and IP3-dependent calcium signals have recently been shown to be required for one or more aspects of *C. elegans* development, we took the set of proteins encoded by this genome (at the time this search was initiated WORMPEP14 was the available
15 predicted protein set) and began searching for proteins which fit the criteria above. This search began by proceeding in alphabetical order through WORMPEP14 and arbitrarily excluding all proteins below approximately 1000 amino acids in size, followed by focusing on
20 remaining proteins with clear TM spanning regions similar to those of other calcium channels. We stopped this screen on encountering a protein designated C05C12.3, a predicted protein of 1816 amino acids (SEQ ID NO:13). C05C12.3 was notable because its central pore region had some sequence similarity to but was clearly distinct from members of the Trp
25 family of calcium channels, and the hydrophobicity plot of this region showed a characteristically wide spacing between the fifth and sixth TM regions for the amino acid residues which are thought to line the channel pore region and mediate the calcium selectivity of the channels. In addition, it lacked any ankyrin repeats in the region amino-terminal to its pore region, further distinguishing it from other Trp family proteins.

30 We then used C05C12.3 for BLAST alignment screening of the rest of the *C. elegans* genome and also mammalian databases for homologous proteins, revealing two other *C. elegans* homologues (SEQ ID NO:14 and SEQ ID NO:15), and also a recently cloned mammalian protein named melastatin-1 (MLSN-1/SOC-1, SEQ ID NOs:9 and 10, and

GenBank Acc. No. AF071787). Using these sequences, we subsequently performed an exhaustive screening of publicly accessible EST databases in search of lymphocyte homologues, but were unsuccessful in detecting any homologous transcripts in any lymphocyte lines. Since MLSN-1 (SEQ ID NOs:9 and 10) was expressed exclusively in melanocytes and retina by Northern blot hybridization and by EST database searching, there was no evidence that this type of channel was expressed in the type of cell in which ICRA-
like currents were best defined. Subsequent BLAST searches picked up mouse EST sequence AI098310 (SEQ ID NO:22) from a monocyte cell line. The I.M.A.G.E. consortium clone containing the above-identified EST was then purchased from ATCC (clone ID. 1312756, Manassas, VA) and was further characterized. Using other portions of this sequence in EST searches, we subsequently picked up similar sequences in human B-cells (SEQ ID NOs:20 and 21), and other cell types as well (SEQ ID NOs: 11, 12, 16, 17, 18, and 19). Most of these sequences were subsequently identified to be part of the 3'-UTR or of the carboxy terminal region of the proteins, which are not readily identifiable as Trp channels, providing an explanation for the art's inability to detect any type of Trp related transcripts in lymphocytes. Partial sequences from the 5' and/or 3' ends of the above identified clones were then used to screen leukocyte and kidney cDNA libraries to extend the original sequences more toward the 5' and/or 3' ends.

In view of the foregoing, it was concluded that channels of this type were expressed in many types of lymphocytes, and therefore were members of a new family of SOC/CRAC calcium channels.

Experimental Procedures

Screening of the cDNA libraries

Leukocyte and kidney cDNA libraries from Life Technologies (Gaithersburg, MD) were screened using the Gene Trapper II methodology (Life Technologies) according to manufacturer's recommendation, using the inserts of I.M.A.G.E. clone ID nos. 1312756 and 1076485 from ATCC (Manassas, VA), under stringent hybridization conditions. Using standard methodology (*Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York), individual cDNA clones were subjected to 3-4 rounds of amplification and purification under the same hybridization conditions.

After excision from the vector and subcloning of inserts into the plasmid forms, several clones were sequenced by the Beth Israel Deaconess Medical Center's Automated

Sequencing Facility. Molecular biological techniques such as restriction enzyme treatment, subcloning, DNA extraction, bacterial culture and purification of DNA fragments were performed according to methods well known in the art. Computer analyses of protein and DNA sequences was done using "Assemblylign" (Oxford Molecular, Cambell, CA). Multiple
5 alignments of the SOC/CRAC family members were produced using the CLUSTAL facility of the MacVector program. Restriction endonucleases, expression vectors, and modifying enzymes were purchased from commercial sources (Gibco-BRL). Sequencing vectors for DNA were purchased from Stratagene (La Jolla, CA).

Once the first members of what appeared to be a novel family of calcium channel
10 receptors were identified and characterized, additional BLAST alignments were performed with the newly characterized nucleic acid sequences. An initial match was with genomic DNA fragment NH0332L11 (Genbank Acc. No. AC005538). Using this genomic sequence, promoters were designed and a number of cDNA libraries was surveyed by PCR. A prostate specific message was identified and characterized, leading to the isolation and
15 characterization of SOC-4/CRAC-3 (SEQ ID NOs: 31 and 32).

Functional Assays

Transient Expression of SOC/CRAC

In our initial transient expression experiments, we expressed or expect to express a SOC/CRAC molecule transiently in RBL-2H3 mast cells, Jurkat T cells, and A20
20 B-lymphocytes using both electroporation and vaccinia virus-driven expression, and measured the calcium influx produced by depletion of intracellular calcium stores with thapsigargin. Each of the foregoing techniques is well known to those of ordinary skill in the art and can be performed using various methods (see, e.g., Current Methods in Molecular Biology, eds. Ausubal, F.M., et al. 1987, Green Publishers and Wiley Interscience, N.Y.,
25 N.Y.). Exemplary methods are described herein.

Depletion of intracellular calcium stores is accomplished by treating the cells with 1 micromolar thapsigargin; alternative agents which function to deplete intracellular stores are described in by Putney, J.W. Jr., in Capacitative Calcium Entry, R.G. Landes Co. and Chapman & Hall, 1997 and include, for example, ionomycin, cyclopiazonic acid, and DBHQ.

30 Calcium influx is determined by measuring cytoplasmic calcium as indicated using the fura-2 fluorescent calcium indicator (see, e.g., G. Grynkiewicz, M. Poenie, R. Y. Tsien, A new generation of Ca²⁺ indicators with greatly improved fluorescence properties, J. Biol

Chem 260, 3440-50 (1985), and M. Poenie, R. Tsien, Fura-2: a powerful new tool for measuring and imaging $[Ca^{2+}]_i$ in single cells, Prog Clin Biol Res 210, 53-6 (1986)).

Patch Clamp Analysis and Determining Selectivity of SOC/CRAC

Patch clamp analysis of cells injected with SOC/CRAC cRNA is performed by using the general patch technique as described in Neher, E., "Ion channels for communication between and within cells", Science, 1992; 256:498-502. Specific techniques for applying the patch clamp analysis to RBL cells are described in Hoth, M., and Penner, R., "Depletion of intracellular calcium stores activates a calcium current in mast cells", Nature, 1992; 355:3535-355. Additional protocols for applying the patch clamp technique to other cell types are described in Putney, J.W. Jr., in Capacitative Calcium Entry, R.G. Landes Co. and Chapman & Hall, 1997

An exemplary protocol for patch clamp analysis of SOC/CRAC molecule expressed in RBL-2H3 mast cells using a recombinant vaccinia virus is as follows. The currents elicited by store depletion are determined using the whole cell configuration (Neher, E., Science, 1992; 256:498-502). Currents in SOC/CRAC expressing cells are compared to currents in control cells expressing an irrelevant protein or a classic Trp family calcium channel known as VR1 (M. J. Caterina, et al., The capsaicin receptor: a heat-activated ion channel in the pain pathway [see comments], Nature 389, 816-24 (1997)) in order to assess the contribution of SOC/CRAC expression. In addition, the magnitude of whole cell currents in the presence of extracellular calcium (10 mM), barium (10 mM), or magnesium (10 mM) are compared to determine the relative permeability of the channels to each of these ions (Hoth, M., and Penner, R., Nature, 1992; 355:3535-355) and, thereby, determine the ionic selectivity.

Pharmacologic Behavior of SOC/CRAC

For analysis of the pharmacologic behavior of a SOC/CRAC molecule, a SOC/CRAC molecule is expressed in RBL-2H3 mast cells using a recombinant vaccinia virus, and the degree of calcium influx elicited by store depletion is monitored using a bulk spectrofluorimeter or a fluorescence microscope and the calcium sensitive dye fura-2 (G. Grynkiewicz, M. Poenie, R. Y. Tsien, A new generation of Ca^{2+} indicators with greatly improved fluorescence properties, J Biol Chem 260, 3440-50 (1985) and M. Poenie, R. Tsien, Fura-2: a powerful new tool for measuring and imaging $[Ca^{2+}]_i$ in single cells, Prog Clin Biol Res 210, 53-6 (1986)). The level of cytoplasmic calcium in SOC/CRAC expressing cells is compared to the level achieved in control cells expressing an irrelevant protein or a classic Trp. family calcium channels known as VR1 (M. J. Caterina, et al., The

-46-

capsaicin receptor: a heat-activated ion channel in the pain pathway [see comments], Nature 389, 816-24 (1997)). These cells then are pre-incubated with the desired pharmacologic reagent, and again the response to store depletion is monitored. Comparison of the effect of depleting stores in SOC/CRAC expressing cells relative to controls in the presence or absence of the pharmacologic reagent is used to assess the ability of that reagent to modulate SOC/CRAC activity. Sphingosine is an exemplary molecule that can be used as pharmacologic reagents for pharmacologic characterization of SOC/CRAC calcium channels. See, e.g., Mathes, C., et al., Calcium release activated calcium current as a direct target for sphingosine, J Biol Chem 273(39):25020-25030 (1998). Other non-specific calcium channel inhibitors that can be used for this purpose include SKR96365 (Calbiochem) and Lanthanum.

Bulk Calcium Assays

Bulk calcium assays can be performed in a PTI Deltascan bulk spectrofluorometer using fura-2 as described in Scharenberg AM, et al., EMBO J, 1995, 14(14):3385-94.

Gene Targeting

The method (and reagents) described by Buerstedde JM et al, (Cell, 1991, Oct 4;67(1):179-88), was used to generate "knockouts" in cells. Briefly, part of the chicken SOC-2/CRAC-1 genomic sequence coding for the transmembrane region was cloned utilizing the human sequence as the probe in a chicken library screen. Chicken SOC-2/CRAC-1 clones were isolated and characterized using standard methodology. The putative exon and domain arrangement of the chicken SOC-2/CRAC-1, is depicted in Figure 1. The exons coding for TM5 (pore region) and TM6, were replaced with promoter/antibiotic cassettes (see Figure1). These targeting vectors were then used to target (and replace) the endogenous gene in DT-40 cells (chicken B lymphocyte cells).

Results

Example 1: Transient Expression of SOC/CRAC

In the above-identified cell lines and using both of the foregoing expression techniques, SOC/CRAC expression enhances thapsigargin-dependent influx. In addition, SOC/CRAC expression also enhances the amount of intracellular calcium stores. That this effect is likely due to SOC/CRAC acting as a plasma membrane calcium channel can be confirmed by producing an in-frame carboxy-terminal translational fusion with green fluorescent protein followed by confocal microscopy, revealing that SOC/CRAC is expressed predominantly as a plasma membrane calcium channel.

Example 2: Patch Clamp Analysis

The biophysical characteristics of SOC/CRAC enhanced currents when expressed in *Xenopus* oocytes are determined. SOC/CRAC cRNA injection is able to enhance thapsigargin-dependent whole cell currents. In addition, SOC/CRAC does not alter the reversal potential of these currents and the determination of the P_{Ca}/P_{Na} ratio shows that SOC/CRAC channels are highly calcium selective.

Example 3: *Pharmacologic Behavior of SOC/CRAC*

The pharmacologic behavior of SOC/CRAC is evaluated as described above. SOC/CRAC-enhanced influx is inhibited by sphingosine in a manner that is substantially the same as that of endogenous thapsigargin-dependent calcium influx.

Example 4: *Gene targeting*

Transfection of DT-40 cells with the foregoing targeting vectors, selection for antibiotic resistance, and screening, is collectively referred to, herein, as a round of targeting. For the first round of targeting SOC-2/CRAC-1, 18/24 clones with homologous recombination of the targeting construct into one of the endogenous SOC-2/CRAC-1 alleles were obtained. On the second round of targeting (in order to target the second allele and therefore generate a homozygous SOC-2/CRAC-1 mutant cell), 0/48 clones were obtained. These results indicate that a "null" SOC-2/CRAC-1 mutation is detrimental to DT-40 cells, and that SOC-2/CRAC-1 is required for cell viability.

Table I. Nucleotide Sequences with homologies to SOC/CRAC nucleic acids

Sequences with SEQ ID NOs and GenBank accession numbers:
SEQ ID NO:9, AB001535, AI226731, H18835, AA419592, AA261842, AA419407, AA592910, D86107, AI098310, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853.

Table II. Amino Acid Sequences with homologies to SOC/CRAC polypeptides

Sequences with SEQ ID NOs and GenBank accession numbers:
SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572.

All references, patents, and patent documents disclosed herein are incorporated by reference herein in their entirety.

What is claimed is presented below and is followed by a Sequence Listing. We claim:

-48-
Claims

1. An isolated nucleic acid molecule, comprising:

(a) nucleic acid molecules which hybridize under stringent conditions to a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;

(b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;

(c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and

(d) complements of (a), (b) or (c).

2. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:1.

3. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:27.

4. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:29.

5. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:31.

6. An isolated nucleic acid molecule selected from the group consisting of

(a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:29, and SEQ ID NO:31,

(b) complements of (a),
provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to any sequence selected from a sequence group consisting of

(1) sequences having the SEQ. ID NOS. or GenBank accession numbers of Table I,

(2) complements of (1), and

(3) fragments of (1) and (2).

7. The isolated nucleic acid molecule of claim 6, wherein the sequence of contiguous nucleotides is selected from the group consisting of:

- (1) at least two contiguous nucleotides nonidentical to the sequence group,
- (2) at least three contiguous nucleotides nonidentical to the sequence group,
- (3) at least four contiguous nucleotides nonidentical to the sequence group,
- (4) at least five contiguous nucleotides nonidentical to the sequence group,
- (5) at least six contiguous nucleotides nonidentical to the sequence group,
- (6) at least seven contiguous nucleotides nonidentical to the sequence group.

8. The isolated nucleic acid molecule of claim 6, wherein the unique fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20 nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, and 200 nucleotides.

9. The isolated nucleic acid molecule of claim 6, wherein the molecule encodes a polypeptide which is immunogenic.

10. An expression vector comprising the isolated nucleic acid molecule of claims 1, 2, 3, 4, 5, 6, 7, 8, or 9 operably linked to a promoter.

11. A host cell transformed or transfected with the expression vector of claim 10.

12. An isolated polypeptide encoded by the isolated nucleic acid molecule according to anyone of claims 1 or 6, wherein the polypeptide comprises a SOC/CRAC polypeptide or a unique fragment thereof.

13. The isolated polypeptide of claim 12, wherein the isolated polypeptide is encoded by the isolated nucleic acid molecule of claim 2, 3, 4, or 5.

14. The isolated polypeptide of claim 13, wherein the isolated polypeptide comprises a polypeptide having the sequence of amino acids selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32.

15. An isolated polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3, 4, or 5, wherein the polypeptide, or unique fragment thereof is immunogenic.

16. An isolated binding polypeptide which binds selectively to a polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3, 4, or 5.

17. The isolated binding polypeptide of claim 16, wherein the isolated binding polypeptide binds to a polypeptide having the sequence of amino acids selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32.

18. The isolated binding polypeptide of claim 17, wherein the isolated binding polypeptide is an antibody or an antibody fragment selected from the group consisting of a Fab fragment, a F(ab)₂ fragment or a fragment including a CDR3 region selective for the polypeptide.

19. An isolated polypeptide, comprising a unique fragment of the polypeptide of claim 12 of sufficient length to represent a sequence unique within the human genome, provided that the fragment excludes a sequence of contiguous amino acids identified in Table II.

20. A method for isolating a SOC/CRAC molecule having SOC/CRAC calcium channel activity, comprising:

a) contacting a binding molecule that is a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing one or more SOC/CRAC molecules, under conditions sufficient to form a complex of the SOC/CRAC nucleic acid or the SOC/CRAC binding polypeptide and the SOC/CRAC molecule;

b) detecting the presence of the complex;

c) isolating the SOC/CRAC molecule from the complex; and

d) determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity.

21. The method of claim 20, wherein the binding molecule is a SOC/CRAC nucleic acid.

22. The method of claim 20, wherein the binding molecule is a SOC/CRAC binding polypeptide.

23. The method of claim 21, wherein the SOC/CRAC nucleic acid comprises at least 14 nucleotides from any contiguous portion of a sequence of nucleotides selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31.

5 24. A method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity, comprising:

a) contacting a SOC/CRAC polypeptide with a candidate agent suspected of modulating SOC/CRAC calcium channel activity, under conditions sufficient to allow the SOC/CRAC polypeptide to interact selectively with the candidate agent;

10 b) detecting a Ca^{2+} concentration associated with SOC/CRAC calcium channel activity of the SOC/CRAC polypeptide in the presence of the candidate agent; and

c) comparing the Ca^{2+} concentration of step (b) with a control Ca^{2+} concentration of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates SOC/CRAC calcium channel activity.

15 25. A method for determining the level of SOC/CRAC expression in a subject, comprising:

a) measuring the expression of SOC/CRAC in a test sample obtained from the subject, and

20 b) comparing the measured expression of SOC/CRAC in the test sample to the expression of the SOC/CRAC polypeptide in a control to determine the level of SOC/CRAC expression in the subject.

25 26. The method of claim 25, wherein the expression of SOC/CRAC in (b) is SOC/CRAC mRNA expression.

27. The method of claim 25, wherein the expression of SOC/CRAC in (b) is SOC/CRAC polypeptide expression.

28. The method of claim 25, wherein the test sample is tissue.

29. The method of claim 25, wherein the test sample is a biological fluid.

30. The method of claim 26, wherein SOC/CRAC mRNA expression is measured using the Polymerase Chain Reaction (PCR).

31. The method of claim 26, wherein SOC/CRAC mRNA expression is measured using a method selected from the group consisting of northern blotting, monoclonal antisera to
5 SOC/CRAC and polyclonal antisera to SOC/CRAC.

32. A kit, comprising a package containing:

an agent that selectively binds to the isolated nucleic acid of claim 1 or an expression product thereof, and

10 a control for comparing to a measured value of binding of said agent to said isolated nucleic acid of claim 1 or expression product thereof.

33. The kit of claim 32, wherein the control comprises an epitope of the expression product of the nucleic acid of claim 1.

34. A pharmaceutical composition comprising:

15 a pharmaceutically effective amount of an agent comprising of an isolated nucleic acid molecule of claim 1 or an expression product thereof, and

a pharmaceutically acceptable carrier.

35. The pharmaceutical composition of claim 34, wherein the agent is an expression product of the isolated nucleic acid molecule of claim 1.

20 36. A method for identifying agents useful in the modulation of a SOC/CRAC polypeptide kinase activity, comprising:

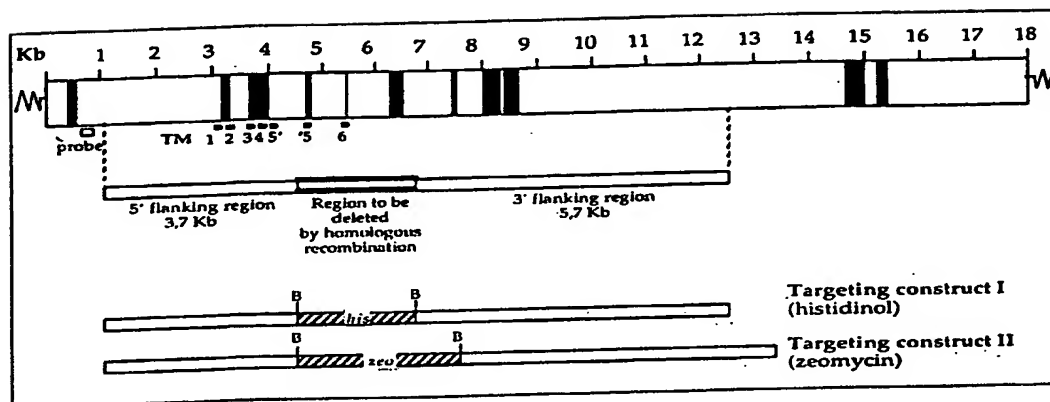
a) contacting a SOC/CRAC polypeptide with kinase activity with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the
25 candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity;

b) detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and

c) comparing the kinase activity of step (b) with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the
30 candidate agent modulates SOC/CRAC kinase activity.

37. The method of claim 36, wherein the SOC/CRAC polypeptide comprises amino acids 999-1180 of the sequence represented as SEQ ID NO:24, or a fragment thereof that retains the kinase activity.

FIGURE 1.



-1-

SEQUENCE LISTING

<110> Beth Israel Deaconess Medical Center, Inc.
Scharenberg, Andrew

<120> CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY

<130> B0662/7026WO/ERP/KA

<150> U.S. 60/114,220

<151> 1998-12-30

<150> U.S. 60/120,018

<151> 1999-01-29

<150> U.S. 60/140,415

<151> 1999-06-22

<160> 32

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 1212

<212> DNA

<213> Homo Sapiens

<400> 1

gcacgaggca aattttttgt tagtacacca tctcagccaa gttgcaaaag ccacttggaa	60
actggaacca aagatcaaga aactgtttgc tctaaagcta cagaaggaga taatacagaa	120
tttgagcat ttgtaggaca cagagatagc atggatttac agaggtttaa agaaacatca	180
aacaagataa aaatactatc caataacaat acttctgaaa acactttgaa acgagtgagt	240
tctcttgctg gatttactga ctgtcacaga acttccattc ctgttcattc aaaacaagaa	300
aaaatcagta gaaggccatc taccgaagac actcatgaag tagattccaa agcagcttta	360
ataccggttt gtagatttca actaaacaga tatatattat taaatacatt aaactttttt	420
agataagatc tacaaagtgg tgatatttgg gactatatca aaaattcaaa aaaatttttc	480
ttaagaaaac tgacttttagc atagtagcag ttacagaaaa gtttcttaca gtgaatagtc	540
aggaatttta aagaaaaatt tatgcagaat aaaggcagga atctcttttt gtttgaattg	600
aagctaatta tatgaactca ttccagcta actgcgataa tgattgattt tgcaaatcc	660
ctttaaaagc acacactgac aagacaaaaa gctcaggaaa aggcagaaaa attactcctt	720
tataatcaag tattatatat aagtcagtgc tcataatttt gctcaagaaa atattgactt	780
acattcatat atatctgttc tggcatagag agattatgtt gttaaaaatca tgttattgaa	840
aaaagtattt tcagtgggga aagagggttag ttaacaaaga gattcacagt aacaaatcct	900
cctttctgga gggactcttc ctgaccctga gctgcacaac tttgcaacaa attaaagcct	960
aaccgaagat gacctcaciaa tggcaattta gaactcatgg gagtcaactt acataaacgg	1020
tatttgattt ctgataagat agtggaaatta ttggttatag atgacaaaaat aagtatgttt	1080
aaagtgatga tggacataaa aaagttttta atataaaaaca tgagaaaaaga aggagatact	1140
attcaaaaag actggcaaat ttgaaaaact agaaataaaa aaaaaaaaaa aaaatgagcg	1200
gccgcaagct tt	1212

<210> 2

<211> 141

<212> PRT

<213> Homo Sapiens

<400> 2

Ala	Arg	Gly	Lys	Phe	Phe	Val	Ser	Thr	Pro	Ser	Gln	Pro	Ser	Cys	Lys
1				5					10					15	
Ser	His	Leu	Glu	Thr	Gly	Thr	Lys	Asp	Gln	Glu	Thr	Val	Cys	Ser	Lys
			20					25					30		

-2-

Ala	Thr	Glu	Gly	Asp	Asn	Thr	Glu	Phe	Gly	Ala	Phe	Val	Gly	His	Arg
	35						40					45			
Asp	Ser	Met	Asp	Leu	Gln	Arg	Phe	Lys	Glu	Thr	Ser	Asn	Lys	Ile	Lys
	50					55					60				
Ile	Leu	Ser	Asn	Asn	Asn	Thr	Ser	Glu	Asn	Thr	Leu	Lys	Arg	Val	Ser
65					70					75					80
Ser	Leu	Ala	Gly	Phe	Thr	Asp	Cys	His	Arg	Thr	Ser	Ile	Pro	Val	His
			85						90					95	
Ser	Lys	Gln	Glu	Lys	Ile	Ser	Arg	Arg	Pro	Ser	Thr	Glu	Asp	Thr	His
		100						105					110		
Glu	Val	Asp	Ser	Lys	Ala	Ala	Leu	Ile	Pro	Val	Cys	Arg	Phe	Gln	Leu
		115					120					125			
Asn	Arg	Tyr	Ile	Leu	Leu	Asn	Thr	Leu	Asn	Phe	Phe	Arg			
	130					135					140				

<210> 3
 <211> 739
 <212> DNA
 <213> Homo Sapiens

<220>
 <221> unsure
 <222> (5)...(5)
 <223> UNKNOWN

<221> unsure
 <222> (21)...(22)
 <223> UNKNOWN

<221> unsure
 <222> (29)...(29)
 <223> UNKNOWN

<400> 3	
tcgantaggg gctctccacc nncatactng gatgatgggt ggtgaagtct atgcatacga	60
aattgatgtg tgtgcaaacg attctgttat ccctcaaatc tgtggtcctg ggacgtgggt	120
gactccattt cttcaagcag tctacctctt tgwacagtat atcattatgg ttaatcttct	180
tattgcattt ytcaacaatg tgtatttaca agtgaaggca atttccaata ttgyatggaa	240
gtaccagcgt tatcatTTta ttatggctta tcatgagaaa ccagttctgc ctcctccact	300
tatcattctt agccatatag tttctctgtt ttgctgcata tgtaagagaa gaaagaaaga	360
taagacttcc gatggaccaa aacttttctt aacagaagaa gatcaaaaga aacttcatga	420
ttttgaagag cagtgtgttg aaatgtattt caatgaaaaa gatgacaaat ttcattctgg	480
gagtgaagag agaattcgtg tcacttttga aagagtggaa cagatgtgca ttcagattaa	540
agaagttgga gatccgtgtc aactacataa aaagatcatt acaatcatta gattctcaaa	600
ttggccattt gcaagatctt tcagccctga cggtagatac attaaaaaca ctactggcc	660
aaaagcgtcg gaagctagca aagttcataa tgaaatcaca cgagaactga gcatttccaa	720
acacttggct caaaacctt	739

<210> 4
 <211> 235
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> UNSURE
 <222> (41)...(41)
 <223> UNKNOWN

<221> UNSURE
 <222> (54)...(54)

-3-

<223> UNKNOWN

<221> UNSURE

<222> (68)...(68)

<223> UNKNOWN

<400> 4

Met	Met	Val	Gly	Glu	Val	Tyr	Ala	Tyr	Glu	Ile	Asp	Val	Cys	Ala	Asn
1				5					10					15	
Asp	Ser	Val	Ile	Pro	Gln	Ile	Cys	Gly	Pro	Gly	Thr	Trp	Leu	Thr	Pro
			20					25					30		
Phe	Leu	Gln	Ala	Val	Tyr	Leu	Phe	Xaa	Gln	Tyr	Ile	Ile	Met	Val	Asn
		35					40					45			
Leu	Leu	Ile	Ala	Phe	Xaa	Asn	Asn	Val	Tyr	Leu	Gln	Val	Lys	Ala	Ile
	50					55					60				
Ser	Asn	Ile	Xaa	Trp	Lys	Tyr	Gln	Arg	Tyr	His	Phe	Ile	Met	Ala	Tyr
65					70					75				80	
His	Glu	Lys	Pro	Val	Leu	Pro	Pro	Pro	Leu	Ile	Ile	Leu	Ser	His	Ile
				85					90					95	
Val	Ser	Leu	Phe	Cys	Cys	Ile	Cys	Lys	Arg	Arg	Lys	Lys	Asp	Lys	Thr
			100					105					110		
Ser	Asp	Gly	Pro	Lys	Leu	Phe	Leu	Thr	Glu	Glu	Asp	Gln	Lys	Lys	Leu
		115					120					125			
His	Asp	Phe	Glu	Glu	Gln	Cys	Val	Glu	Met	Tyr	Phe	Asn	Glu	Lys	Asp
	130					135					140				
Asp	Lys	Phe	His	Ser	Gly	Ser	Glu	Glu	Arg	Ile	Arg	Val	Thr	Phe	Glu
145					150					155					160
Arg	Val	Glu	Gln	Met	Cys	Ile	Gln	Ile	Lys	Glu	Val	Gly	Asp	Pro	Cys
			165					170						175	
Gln	Leu	His	Lys	Lys	Ile	Ile	Thr	Ile	Ile	Arg	Phe	Ser	Asn	Trp	Pro
			180					185					190		
Phe	Ala	Arg	Ser	Phe	Ser	Pro	Asp	Gly	Arg	Tyr	Ile	Lys	Asn	Thr	His
		195					200					205			
Trp	Pro	Lys	Ala	Ser	Glu	Ala	Ser	Lys	Val	His	Asn	Glu	Ile	Thr	Arg
	210					215					220				
Glu	Leu	Ser	Ile	Ser	Lys	His	Leu	Ala	Gln	Asn					
225					230					235					

<210> 5

<211> 1579

<212> DNA

<213> Homo Sapiens

<220>

<221> unsure

<222> (368)...(368)

<223> g or c

<221> unsure

<222> (372)...(372)

<223> g or c

<221> unsure

<222> (374)...(374)

<223> g or a

<221> unsure

<222> (375)...(375)

<223> g or c

-4-

<221> unsure
<222> (387)... (387)

<221> unsure
<222> (482)... (482)

<400> 5

acgtgcgctg	caggtaccgg	tccggaattc	ccgggtcgac	ccacgcgtcc	ggcatgggtgt	60
tgtaaatata	cttagctcct	ctcttcctca	aggtagatctt	gaaagtaata	atcctttttca	120
ttgtaatat	ttaatgaaag	atgacaaaga	tccccagtg	aatatatattg	gtcaagactt	180
acctgcagta	ccccagagaa	aagaatttaa	ttttccagag	gctgggttctt	cttctgggtgc	240
cttattccca	agtgtgtgtt	ccctccaga	actgcgacag	agactacatg	gggtagaact	300
cttaaaaaata	tttaataaaa	atcaaaaatt	aggcagttca	tctactagca	taccacatct	360
gtcatccsca	csarscaaat	tttttgntag	tacaccatct	cagccaagtt	gcaaaaagcca	420
cttggaact	ggaaccaaaag	atcaagaaac	tgtttgctct	aaagctacag	aaggagataa	480
tncagaattt	ggagcatttg	taggacacag	agatagcatg	gatttacaga	ggtttaaaga	540
aacatcaaac	aagataaaaa	tactatccaa	taacaatact	tctgaaaaca	cttgaaaacg	600
agtgtgttct	cttgcctggat	ttactgactg	tcacagaact	tccattcctg	ttcattcaaa	660
acaagaaaaa	atcagtagaa	ggccatctac	cgaagacact	catgaagtag	attccaaagc	720
agctttaata	ccgttttgta	gatttcaact	aaacagatat	atattattaa	atacatataa	780
ctttttttaga	taagatctac	aaagtgggtga	tatttgggac	tatatcaaaa	attcaaaaaa	840
atttttctta	agaaaactga	ctttagcata	gtagcagtta	cagaaaagtt	tcttacagt	900
aatagtcagg	aatttttaaag	aaaaatttat	gcagaataaa	ggcaggaatc	tctttttgtt	960
tgaattgaag	ctaattatat	gaactcattt	ccagctaact	gcgataatga	ttgattttgc	1020
aaattccctt	taaaagcaca	cactgacaag	acaaaaagct	caggaaaagg	cagaaaaatt	1080
actcctttat	aatcaagtat	tatatataag	tcagtgtctca	taattttgct	caagaaaata	1140
ttgacttaca	ttcatatata	tctgttctgg	catagagaga	ttatgttgtt	aaaatcatgt	1200
tattgaaaaa	agttatttca	gtggggaaag	aggtagtga	acaaagagat	tcacagtaac	1260
aaatcctcct	ttctggagg	actcttctctg	accctgagct	gcacaacttt	gcaacaaatt	1320
aaagcctaac	cgaagatgac	ctcacaatgg	caatttagaa	ctcatgggag	tcaacttaca	1380
taaacgggtat	ttgatttctg	ataagatagt	ggaattattg	gttatagatg	acaaaaataag	1440
tatgttttaa	gtgatgatgg	acataaaaaa	gttttaaata	taaaacatga	gaaaagaagg	1500
agatactatt	caaaaagact	ggcaaatgtg	aaaaactaga	aataaaaaaa	aaaaaaaaaa	1560
atgagcggcc	gcaagcttt					1579

<210> 6
<211> 243
<212> PRT
<213> Homo Sapiens

<220>
<221> UNSURE
<222> (103)... (105)
<223> UNKNOWN

<221> UNSURE
<222> (109)... (109)
<223> UNKNOWN

<221> UNSURE
<222> (141)... (141)
<223> UNKNOWN

<400> 6

Val	Asn	Thr	Leu	Ser	Ser	Ser	Leu	Pro	Gln	Gly	Asp	Leu	Glu	Ser	Asn
1				5					10					15	
Asn	Pro	Phe	His	Cys	Asn	Ile	Leu	Met	Lys	Asp	Asp	Lys	Asp	Pro	Gln
			20					25					30		
Cys	Asn	Ile	Phe	Gly	Gln	Asp	Leu	Pro	Ala	Val	Pro	Gln	Arg	Lys	Glu
	35						40					45			

-5-

Phe	Asn	Phe	Pro	Glu	Ala	Gly	Ser	Ser	Ser	Gly	Ala	Leu	Phe	Pro	Ser
50						55					60				
Ala	Val	Ser	Pro	Pro	Glu	Leu	Arg	Gln	Arg	Leu	His	Gly	Val	Glu	Leu
65					70					75					80
Leu	Lys	Ile	Phe	Asn	Lys	Asn	Gln	Lys	Leu	Gly	Ser	Ser	Ser	Thr	Ser
				85					90					95	
Ile	Pro	His	Leu	Ser	Ser	Xaa	Xaa	Xaa	Lys	Phe	Phe	Xaa	Ser	Thr	Pro
			100					105					110		
Ser	Gln	Pro	Ser	Cys	Lys	Ser	His	Leu	Glu	Thr	Gly	Thr	Lys	Asp	Gln
		115				120						125			
Glu	Thr	Val	Cys	Ser	Lys	Ala	Thr	Glu	Gly	Asp	Asn	Xaa	Glu	Phe	Gly
	130					135					140				
Ala	Phe	Val	Gly	His	Arg	Asp	Ser	Met	Asp	Leu	Gln	Arg	Phe	Lys	Glu
145					150					155					160
Thr	Ser	Asn	Lys	Ile	Lys	Ile	Leu	Ser	Asn	Asn	Asn	Thr	Ser	Glu	Asn
			165					170						175	
Thr	Leu	Lys	Arg	Val	Ser	Ser	Leu	Ala	Gly	Phe	Thr	Asp	Cys	His	Arg
			180				185						190		
Thr	Ser	Ile	Pro	Val	His	Ser	Lys	Gln	Glu	Lys	Ile	Ser	Arg	Arg	Pro
	195					200						205			
Ser	Thr	Glu	Asp	Thr	His	Glu	Val	Asp	Ser	Lys	Ala	Ala	Leu	Ile	Pro
	210					215					220				
Val	Cys	Arg	Phe	Gln	Leu	Asn	Arg	Tyr	Ile	Leu	Leu	Asn	Thr	Leu	Asn
225					230					235					240
Phe	Phe	Arg													

<210> 7
 <211> 3532
 <212> DNA
 <213> Mus Musculus

<220>
 <221> unsure
 <222> (2420)... (2420)
 <223> unknown

<221> unsure
 <222> (2434)... (2434)
 <223> unknown

<221> unsure
 <222> (2461)... (2461)
 <223> unknown

<221> unsure
 <222> (2466)... (2466)
 <223> unknown

<221> unsure
 <222> (2470)... (2470)
 <223> unknown

<400> 7

attatggcctt	atcatgaaaa	accagtcctg	cctcctcctc	ttatcatcct	cagccatata	60
gtttcactgt	tttgcgtgtg	atgcaaaaag	agaagaaaag	ataagacttc	cgatgggcca	120
aaacttttct	taacagaaga	agatcaaaaag	aaactccatg	attttgaaga	gcagtgtgtt	180
gagatgtact	ttgatgagaa	agatgacaaa	ttcaattctg	ggagtgaaga	gagaatccgg	240
gtcacttttg	aaagagtggg	gcagatgagc	attcagatta	aagaagttgg	agatcgtgtc	300
aactacataa	aaagatcatt	acagtcctta	gatttctcaa	ttgggtcatc	gcaagatctc	360

-6-

tcagccctaa	cagtagatac	attgaaaaca	cttacagccc	agaaagcttc	agaagctagt	420
aaagtgacac	atgagatcac	acgagaattg	agtatttcca	aacacttggc	tcagaatctt	480
attgatgatg	ttcctgtaag	acctttgtgg	gaagaacctt	gtgctgtaaa	cacactgagt	540
tcctctcttc	ctcaagggtga	tcgggaaagt	aataatcctt	ttctttgtaa	tatttttatg	600
aaagatgaaa	aagaccccca	atataatctg	tttggaacag	atttgcccgt	gataccccag	660
agaaaaaat	tcaacattcc	agaggctggt	tcctcctgtg	gtgccttatt	cccaagtgt	720
gtttctcccc	cagaattacg	acagagacga	catggggtag	aaatgttaaa	aatatttaaat	780
aaaaatcaaa	aattaggcag	ttcacctaata	agttcaccac	atatgtcctc	cccaccaacc	840
aaattttctg	tgagtacccc	atccccagcca	agttgcaaaa	gtcacttgga	atccacaacc	900
aaagatcaag	aacccatttt	ctataaagct	gcagaagggg	ataacataga	atttgaggca	960
tttgtgggac	acagagatag	tatggactta	cagaggttta	aagaaacatc	aaacaaaata	1020
agagaactgt	tatctaataga	tactcctgaa	aacactctga	aacatgtggg	tgctgttgga	1080
tataagtgaat	gttgtaagac	ttctacttct	cttctactcg	tgcaagcaga	aagctgtagt	1140
agaagagcgt	cgacggaaga	ctctccagaa	gtcgattcta	aagcagcttt	gttaccggat	1200
tggttacgag	atagaccatc	aaacagagaa	atgccatctg	aaggaggaaac	attaaatggt	1260
cttgcttctc	catttaagcc	cgttttggat	acaaattact	attattcagc	tgtggaaaga	1320
aaataacctga	tgaggtgtgc	acagagtatt	ccctctgttc	ctgtacctcc	acgaggcgag	1380
cctgtcacag	tgtaccgtct	ggaggagagt	tctcccagta	tactgaataa	cagcatgtct	1440
tcattggtctc	agctaggcct	ctgtgccaaa	attgagtttt	taagtaaaga	ggaaatggaa	1500
ggtggtttac	gaagagcagt	caaagtgtctg	tgtacctggg	cagagcacga	tatcctgaag	1560
tcagggcatc	tctatatcat	taagtcattt	cttcttgagg	tgataaacac	atggtcaagc	1620
atttataaag	aagatacggg	tctacatctc	tgtctcagag	aaatacaaca	acagagagca	1680
gcacaaaagc	tcacatttgc	ctttaatcag	atgaaaccca	aatccatacc	atattctcca	1740
aggttccttg	aagttttcct	gttgtagctgc	cattcagcag	ggcagtggtt	tgctgtagaa	1800
gagtgcatga	ctggtgaatt	tagaaaatac	aacaacaata	atggtgatga	aatcattcct	1860
acaaataactc	tagaagagat	catgctagcc	tttagccact	ggacctatga	atataccaga	1920
ggggagttac	tggtacttga	cttacaagga	gtgggagaaa	acttgactga	cccatctgta	1980
ataaaaagctg	aagaaaaaag	atcctgtgac	atggtttttg	gccctgcaa	tctaggagaa	2040
gatgcaataa	aaaacttcaa	gagccaaaaca	tccactgtaa	ttcttgcgtg	cgaagactta	2100
aacttcccag	atttgaagag	gaatgactac	acgcccttga	taaaattata	tttctcagg	2160
atgagtcatc	agatttgaat	cttcaatctg	gaaattccac	caaagaatca	gaagcaacaa	2220
attctgttctg	tctgatgtta	tagtgctgag	tcattgggtt	ttgcctacac	ttcacaaaag	2280
tgtaactgtc	agttttcctt	tcgggggaat	tgatgatata	ggaagatgtg	tgcaaaatga	2340
gcttgcctggc	cccacacata	gtctagaggt	aatgttctca	ttgaaaaacg	cctggagggtg	2400
gaggctgcag	atgccagtgn	aaagtgtctag	ctgncagaga	gtcagtgctc	tcgggctgggt	2460
naaggncggn	acccttgctg	ctgagagtgg	tggttctctt	cacctgggtg	aggaccatta	2520
accaaagtca	agtcttcaga	tttgattggc	tgctcagtc	cagcccatc	agctaaggaa	2580
actaaattgc	gcagcttttt	aaatggctga	agtcttctc	agtttgcgt	ctatgataat	2640
gatgttagct	ctcaactagg	tgtttggg	cacgggagaa	ctactcctta	caattttgct	2700
tcacaggcat	gttacaagc	ctgcactgaa	aaccgtttgt	cttccctctc	tcctccctc	2760
ttttccctgt	agtattgagg	atcaaaacca	gggcctcatg	aagaccattt	tctaagagac	2820
attttattta	agaatcaact	atagagtcta	tgtttatgga	tacagccagt	ttttgttaa	2880
caaaacctga	attgtgcaaa	agggtttttt	aacatttatc	aatgttaagt	aaaagaaagc	2940
catgataaat	agaatttaac	tcaactgttc	atgggtgttt	cctgtgagga	aggttacagt	3000
tgtaacagcc	tgagtttga	tacatctcca	aagatttaca	gacttagtgt	atcaaatcag	3060
agtgtcatgt	gagctctcac	attgaaaatt	ctataggaat	gtgtcaatgt	gaattctatt	3120
tctggtactt	aagaaatcag	ttgttggatt	atccttatac	agtataggga	gatcacaata	3180
caactttatg	ccaataaaat	ctaacttaac	tgcccagata	tttttgcata	tttagcaaca	3240
agaaaagctt	atcatttgac	tcaagtttta	tgctttctct	ttcttttcat	ttcctaggta	3300
ctaattttta	tttttatttg	gaaggagcag	tgtaaaagctt	acttgtattc	aatagtgtat	3360
ctcatagata	cagacaaggc	cgcagagata	agctgtttaa	tagtgtttta	tggtgatgtg	3420
gagagaaagg	tgtattactt	aaaaatacta	taccatatac	gttttgata	tcattaaatc	3480
tttaaaagaa	attaaattta	ttcttgttta	aaaaaaaaaa	aaaaaaaaaa	aa	3532

<210> 8
 <211> 475
 <212> PRT
 <213> Mus Musculus

<400> 8

-7-

Ile	Met	Ala	Tyr	His	Glu	Lys	Pro	Val	Leu	Pro	Pro	Pro	Leu	Ile	Ile
1				5					10				15		
Leu	Ser	His	Ile	Val	Ser	Leu	Phe	Cys	Cys	Val	Cys	Lys	Arg	Arg	Lys
			20					25					30		
Lys	Asp	Lys	Thr	Ser	Asp	Gly	Pro	Lys	Leu	Phe	Leu	Thr	Glu	Glu	Asp
		35					40					45			
Gln	Lys	Lys	Leu	His	Asp	Phe	Glu	Glu	Gln	Cys	Val	Glu	Met	Tyr	Phe
	50					55					60				
Asp	Glu	Lys	Asp	Asp	Lys	Phe	Asn	Ser	Gly	Ser	Glu	Glu	Arg	Ile	Arg
65					70					75					80
Val	Thr	Phe	Glu	Arg	Val	Glu	Gln	Met	Ser	Ile	Gln	Ile	Lys	Glu	Val
				85					90					95	
Gly	Asp	Arg	Val	Asn	Tyr	Ile	Lys	Arg	Ser	Leu	Gln	Ser	Leu	Asp	Ser
			100					105						110	
Gln	Ile	Gly	His	Leu	Gln	Asp	Leu	Ser	Ala	Leu	Thr	Val	Asp	Thr	Leu
		115					120					125			
Lys	Thr	Leu	Thr	Ala	Gln	Lys	Ala	Ser	Glu	Ala	Ser	Lys	Val	His	Asn
	130					135					140				
Glu	Ile	Thr	Arg	Glu	Leu	Ser	Ile	Ser	Lys	His	Leu	Ala	Gln	Asn	Leu
145					150					155					160
Ile	Asp	Asp	Val	Pro	Val	Arg	Pro	Leu	Trp	Glu	Glu	Pro	Ser	Ala	Val
				165					170						175
Asn	Thr	Leu	Ser	Ser	Ser	Leu	Pro	Gln	Gly	Asp	Arg	Glu	Ser	Asn	Asn
			180					185						190	
Pro	Phe	Leu	Cys	Asn	Ile	Phe	Met	Lys	Asp	Glu	Lys	Asp	Pro	Gln	Tyr
		195					200					205			
Asn	Leu	Phe	Gly	Gln	Asp	Leu	Pro	Val	Ile	Pro	Gln	Arg	Lys	Glu	Phe
	210					215					220				
Asn	Ile	Pro	Glu	Ala	Gly	Ser	Ser	Cys	Gly	Ala	Leu	Phe	Pro	Ser	Ala
225					230					235					240
Val	Ser	Pro	Pro	Glu	Leu	Arg	Gln	Arg	Arg	His	Gly	Val	Glu	Met	Leu
				245						250				255	
Lys	Ile	Phe	Asn	Lys	Asn	Gln	Lys	Leu	Gly	Ser	Ser	Pro	Asn	Ser	Ser
			260					265					270		
Pro	His	Met	Ser	Ser	Pro	Pro	Thr	Lys	Phe	Ser	Val	Ser	Thr	Pro	Ser
		275					280					285			
Gln	Pro	Ser	Cys	Lys	Ser	His	Leu	Glu	Ser	Thr	Thr	Lys	Asp	Gln	Glu
	290					295					300				
Pro	Ile	Phe	Tyr	Lys	Ala	Ala	Glu	Gly	Asp	Asn	Ile	Glu	Phe	Gly	Ala
305					310					315					320
Phe	Val	Gly	His	Arg	Asp	Ser	Met	Asp	Leu	Gln	Arg	Phe	Lys	Glu	Thr
			325					330						335	
Ser	Asn	Lys	Ile	Arg	Glu	Leu	Leu	Ser	Asn	Asp	Thr	Pro	Glu	Asn	Thr
			340					345					350		
Leu	Lys	His	Val	Gly	Ala	Ala	Gly	Tyr	Ser	Glu	Cys	Cys	Lys	Thr	Ser
		355					360					365			
Thr	Ser	Leu	His	Ser	Val	Gln	Ala	Glu	Ser	Cys	Ser	Arg	Arg	Ala	Ser
						375					380				
Thr	Glu	Asp	Ser	Pro	Glu	Val	Asp	Ser	Lys	Ala	Ala	Leu	Leu	Pro	Asp
385					390					395					400
Trp	Leu	Arg	Asp	Arg	Pro	Ser	Asn	Arg	Glu	Met	Pro	Ser	Glu	Gly	Gly
				405					410					415	
Thr	Leu	Asn	Gly	Leu	Ala	Ser	Pro	Phe	Lys	Pro	Val	Leu	Asp	Thr	Asn
			420					425					430		
Tyr	Tyr	Tyr	Ser	Ala	Val	Glu	Arg	Asn	Asn	Leu	Met	Arg	Leu	Ser	Gln
		435					440					445			
Ser	Ile	Pro	Phe	Val	Pro	Val	Pro	Pro	Arg	Gly	Glu	Pro	Val	Thr	Val
	450					455					460				
Tyr	Pro	Ser	Gly	Gly	Arg	Val	Leu	Pro	Val	Tyr					
465					470					475					

<210> 9
 <211> 5433
 <212> DNA
 <213> Mus Musculus

<220>
 <221> unsure
 <222> (5094)... (5094)
 <223> unknown

<400> 9
 ggctgaaaga gcctgagctg tgcctctcca ttccactgct gtggcagggt cagaaatctt 60
 gगतगगगग aaccttttgc aaacgggaat gtatctttgt aattcctagc acgaaagact 120
 ctaacagggtg ttgctgtggc cagttcacca accagcatat cccccctctg ccaagtgcac 180
 caccagcaa aaatgaagag gaaagcaaac aggtggagac tcagcctgag aaatgggtctg 240
 ttgccaagca caccagagc tacccaacag attcctatgg agttcttgaa ttccagggtg 300
 gcgatatctt caataaagcc atgtatatcc gtgtatccta tgacaccaag ccagactcac 360
 tgctccatct catggtgaaa gattggcagc tggaaactcc caagctctta atatctgtgc 420
 atggaggcct ccagaacttt gagatgcagc ccaagctgaa acaagtcttt gggaaaggcc 480
 tgatcaaggc tgctatgacc accggggcct ggatcttcac cgggggtgtc agcacagggtg 540
 ttatcagcca cgtaggggat gccttgaaag accactcctc caagtccaga ggccgggttt 600
 gtgctatagg aattgctcca tggggcatcg tggagaataa ggaagacctg gttggaaagg 660
 atgtaacaag agtgtaccag accatgtcca accctctaag taagctctct gtgctcaaca 720
 actccacac ccacttcac ctggctgaca atggcaccct gggcaagtat ggcgcggagg 780
 tgaagctgcy aaggctgctg gaaaagcaca tctccctcca gaagatcaac acaagactgg 840
 ggcaggcggt gccctcgtg ggtctcgtgg tggagggggg ccctaactgt gtgtccatcg 900
 tcttggaata cctgaagaa gagcctccca tccctgtggg gatttgtgat ggcagcggac 960
 gtgcctcgga catcctgtcc tttgcgcaca agtactgtga agaaggcgga ataataaatg 1020
 agtccctcag gtagcagctt ctagttaaca ttcagaaaac atttaattat aataaggcac 1080
 aatcacatca gctgtttgca attataatgg agtgcataaa gaagaaagaa ctgctcactg 1140
 tgttcagaat ggttctgag ggccagcagg acatcgagat ggcaatttta actgccttgc 1200
 tgaaggaac aaacgtatct gtccagatc agctgagctt ggactgggt tggaaaccgg 1260
 tggacatagc acgaagccag atctttgtct ttgggcccc ctggacgccc ctgggaagcc 1320
 tggcaccccc gacggacagc aaagccacgg agaaggagaa gaagccaccc atggccacca 1380
 ccaagggagg aagaggaaaa gggaaaggca agaagaaagg gaaagtgaag gaggaagtgg 1440
 aggaagaaac tgacccccgg aagatagagc tgctgaactg ggtgaatgct ttggagcaag 1500
 cgatgctaga tgcttttagtc ttagatcgtg tcgactttgt gaagctcctg attgaaaacg 1560
 gagtgaacat gcaacacttt ctgaccattc cgaggctgga ggagctctat aacacaagac 1620
 tgggtccacc aaacacactt catctgctgg tgagggatgt gaaaaagagc aaccttccgc 1680
 ctgattacca catcagcctc atagacatcg ggctcgtgct ggagtacctc atgggaggag 1740
 cctaccgctg caactacact cggaaaaact ttcggaccct ttacaacaac ttgtttggac 1800
 caaagaggcc taaagctctt aaacttctgg gaatggaaga tgatgagcct ccagctaaag 1860
 ggaagaaaaa aaaaaaaaag aaaaaggagg aagagatcga cattgatgtg gacgaccctg 1920
 ccgtgagtcg gttccagtat ccctccacg agctgatggg gtgggcagtg ctgatgaaac 1980
 gccagaaaaa ggcagtgttc ctctggcagc gaggggaaga gagcatggcc aaggccctgg 2040
 tggcctgcaa gctctacaag gccatggccc acgagtcctc cgagagtgat ctggtggatg 2100
 acatctccca ggacttggat aacaattcca aagacttcgg ccagcttgct ttggagttat 2160
 tagaccagtc ctataagcat gacgagcaga tcgctatgaa actcctgacc tacgagctga 2220
 aaaactggag caactcgacc tgcctcaaac tggccgtggc agccaaacac cgggacttca 2280
 ttgctcacac ctgcagccag atgctgctga ccgatatgtg gatgggaaga ctgcggatgc 2340
 ggaagaaccc cggcctgaag gttatcatgg ggattcttct acccccacc atcttgtttt 2400
 tggaaatttc cacatatgat gatttctcgt atcaaacatc caaggaaaac gaggatggca 2460
 aagaaaaaga agaggaaat acggatgcaa atgcagatgc tggctcaaga aagggggatg 2520
 aggagaacga gcataaaaaa cagagaagta ttccatcgg acaaaagatc tgtgaattct 2580
 ataacgcgcc cattgtcaag ttctggtttt acacaatatc atacttgggc tacctgctgc 2640
 tcttctaata cgtcatcctg gtgcggatgg atggctggcc gtccctccag gagtggatcg 2700
 tcatctecta catcgtgagc ctggcgtag agaagatag agagatcctc atgtcagaac 2760
 caggcaaact cagccagaaa atcaaagttt ggcttcagga gtactggaac atcacagatc 2820
 tcgtggccat ttccacattc atgattggag caattcttcg cctacagaac cagccctaca 2880

```

tgggctatgg ccgggtgatc tactgtgtgg atatcatctt ctggtacatc cgtgtcctgg 2940
acatcttttg tgtcaacaag tatctggggc catacgtgat gatgattgga aagatgatga 3000
tcgacatgct gtacttttgg gtcacatgct tgggtcgtgct catgagtttc ggagtagccc 3060
gtcaagccat tctgcatcca gaggagaagc cctcttggaa actggcccga aacatcttct 3120
acatgcccta ctggatgatc tatggagagg tgtttgcaga ccagatagac ctctacgcca 3180
tggaaattaa tcctccttgt ggtgagaacc tatatgatga ggagggcaag cggcttcctc 3240
cctgtatccc cggcgccctg ctcactccag cactcatggc gtgctatcta ctggtcgcca 3300
acatcctgct ggtgaacctg ctgattgctg tgttcaacaa tactttcttt gaagtaaaat 3360
caatatccaa ccagggtgtg aagttccagc gatatcagct gattatgaca ttcatgaca 3420
ggccagtcct gccccaccg atgatcattt taagccacat ctacatcacc attatgcgtc 3480
tcagcggccg ctgcaggaaa aagagagaag gggaccaaga ggaacgggat cgtggattga 3540
agctcttcct tagcgacgag gagctaaaga ggctgcatga gttcgaggag cagtgcgtgc 3600
aggagcactt ccgggagaag gaggatgagc agcagtcgtc cagcgacgag cgcatccggg 3660
tcacttctga aagagttgaa aatatgtcaa tgaggttgga agaaatcaat gaaagagaaa 3720
cttttatgaa aacttcctg cagactgttg accttcgact tgctcagcta gaagaattat 3780
ctaacagaat ggtgaatgct ctgaaaaatc ttgcgggaat cgacaggtct gacctgatcc 3840
aggcacggtc ccgggcttct tctgaatgtg aggcaacgta tcttctccg caaagcagca 3900
tcaatagcgc tgatggctac agcttgtatc gatatcattt taacggagaa gagttattat 3960
ttgaggatac atctctctcc acgtcaccag ggacaggagt caggaaaaaa acctgttcct 4020
tccgtataaa ggaagagaag gacgtgaaaa cgcacctagt ccagaatgt cagaacagtc 4080
ttcaccttct actgggcaca agcacatcag caaccccaga tggcagtcac ctgacagtag 4140
atgacttaaa gaacgctgaa gagtcaaaat taggtccaga tattgggatt tcaaaggaaag 4200
atgatgaaag acagacagac tctaaaaaag aagaaactat ttccccaagt ttaataaaaa 4260
cagatgtgat acatggacag gacaaatcag atgttcaaaa cactcagcta acagtggaaa 4320
cgacaaaatc agaaggcact atttcctatc ccctggaaga aacaaaaatt acacgctatt 4380
tccccgatga aactgcaaat gcttgtaaaa caatgaagtc cagaagcttc gtctattccc 4440
ggggaagaaa gctggctcgtt ggggttaacc aggatgtaga gtacagttca atcacggacc 4500
agcaattgac gacggaatgg caatgccaag ttcaaaagat cagcgctct catagcacag 4560
atattcctta cattgtgtcg gaagctgcag tgcaagctga gcaaaaagag cagtttgtag 4620
atatgcaaga tgaacaccat gtcgctgaag caattcctcg aatcctcgc ttgtccctaa 4680
ccattactga cagaaatggg atggaaaact tactgtctgt gaagccagat caaactttgg 4740
gattcccatc tctcaggtca aaaagtttac atggacatcc taggaatgtg aaatccattc 4800
agggaaagtt agacagatct ggacatgcc aatgtgtgta cagcttagta attgtgtctg 4860
gaatgacagc agaagaaaaa aaggttaaga aagagaaagc ttccacagaa actgaatgct 4920
agtctgtttt gtttctttta tttttttttt taacagtcag aaaccacta atgggtgtca 4980
tcttggtcca tctaaacac atmtccaatt tctaaaaaac attttccctt aaaaaatttt 5040
ggaaattcag acttgattta caatttaatg cactaaaagt agtattttgt tagnatatgt 5100
tagtaggctt agttttttca gttgcagtag tatcaaatga aagtgatgat actgtaacga 5160
agataaattg gactaatcag atacaagatt tactaatctc tttattactg agggccacca 5220
aatagcctag gaagtgcctt cgagcactga agtcaccatt aggtcactca agaagtaagc 5280
aactagctgg gcacagtggc tcatgcctgt aatcctagca ctttgggagg ccaaggcaga 5340
aagatagctt gagtccagga gtttgagacc agcctgggca acatagtgat accccatctc 5400
ttaaaaaaaaa aaaaaaaaaa ctgcctcctg gcc 5433

```

<210> 10
 <211> 1533
 <212> PRT
 <213> Mus Musculus

<400> 10
 Met Tyr Ile Arg Val Ser Tyr Asp Thr Lys Pro Asp Ser Leu Leu His
 1 5 10 15
 Leu Met Val Lys Asp Trp Gln Leu Glu Leu Pro Lys Leu Leu Ile Ser
 20 25 30
 Val His Gly Gly Leu Gln Asn Phe Glu Met Gln Pro Lys Leu Lys Gln
 35 40 45
 Val Phe Gly Lys Gly Leu Ile Lys Ala Ala Met Thr Thr Gly Ala Trp
 50 55 60

-10-

Ile	Phe	Thr	Gly	Gly	Val	Ser	Thr	Gly	Val	Ile	Ser	His	Val	Gly	Asp	65	70	75	80
Ala	Leu	Lys	Asp	His	Ser	Ser	Lys	Ser	Arg	Gly	Arg	Val	Cys	Ala	Ile	85	90	95	
Gly	Ile	Ala	Pro	Trp	Gly	Ile	Val	Glu	Asn	Lys	Glu	Asp	Leu	Val	Gly	100	105	110	
Lys	Asp	Val	Thr	Arg	Val	Tyr	Gln	Thr	Met	Ser	Asn	Pro	Leu	Ser	Lys	115	120	125	
Leu	Ser	Val	Leu	Asn	Asn	Ser	His	Thr	His	Phe	Ile	Leu	Ala	Asp	Asn	130	135	140	
Gly	Thr	Leu	Gly	Lys	Tyr	Gly	Ala	Glu	Val	Lys	Leu	Arg	Arg	Leu	Leu	145	150	155	160
Glu	Lys	His	Ile	Ser	Leu	Gln	Lys	Ile	Asn	Thr	Arg	Leu	Gly	Gln	Gly	165	170	175	
Val	Pro	Leu	Val	Gly	Leu	Val	Val	Glu	Gly	Gly	Pro	Asn	Val	Val	Ser	180	185	190	
Ile	Val	Leu	Glu	Tyr	Leu	Gln	Glu	Pro	Pro	Ile	Pro	Val	Val	Ile		195	200	205	
Cys	Asp	Gly	Ser	Gly	Arg	Ala	Ser	Asp	Ile	Leu	Ser	Phe	Ala	His	Lys	210	215	220	
Tyr	Cys	Glu	Glu	Gly	Gly	Ile	Ile	Asn	Glu	Ser	Leu	Arg	Glu	Gln	Leu	225	230	235	240
Leu	Val	Thr	Ile	Gln	Lys	Thr	Phe	Asn	Tyr	Asn	Lys	Ala	Gln	Ser	His	245	250	255	
Gln	Leu	Phe	Ala	Ile	Ile	Met	Glu	Cys	Met	Lys	Lys	Lys	Glu	Leu	Val	260	265	270	
Thr	Val	Phe	Arg	Met	Gly	Ser	Glu	Gly	Gln	Gln	Asp	Ile	Glu	Met	Ala	275	280	285	
Ile	Leu	Thr	Ala	Leu	Leu	Lys	Gly	Thr	Asn	Val	Ser	Ala	Pro	Asp	Gln	290	295	300	
Leu	Ser	Leu	Ala	Leu	Ala	Trp	Asn	Arg	Val	Asp	Ile	Ala	Arg	Ser	Gln	305	310	315	320
Ile	Phe	Val	Phe	Gly	Pro	His	Trp	Thr	Pro	Leu	Gly	Ser	Leu	Ala	Pro	325	330	335	
Pro	Thr	Asp	Ser	Lys	Ala	Thr	Glu	Lys	Glu	Lys	Lys	Pro	Pro	Met	Ala	340	345	350	
Thr	Thr	Lys	Gly	Gly	Arg	Gly	Lys	Gly	Lys	Gly	Lys	Lys	Lys	Gly	Lys	355	360	365	
Val	Lys	Glu	Glu	Val	Glu	Glu	Glu	Thr	Asp	Pro	Arg	Lys	Ile	Glu	Leu	370	375	380	
Leu	Asn	Trp	Val	Asn	Ala	Leu	Glu	Gln	Ala	Met	Leu	Asp	Ala	Leu	Val	385	390	395	400
Leu	Asp	Arg	Val	Asp	Phe	Val	Lys	Leu	Leu	Ile	Glu	Asn	Gly	Val	Asn	405	410	415	
Met	Gln	His	Phe	Leu	Thr	Ile	Pro	Arg	Leu	Glu	Glu	Leu	Tyr	Asn	Thr	420	425	430	
Arg	Leu	Gly	Pro	Pro	Asn	Thr	Leu	His	Leu	Leu	Val	Arg	Asp	Val	Lys	435	440	445	
Lys	Ser	Asn	Leu	Pro	Pro	Asp	Tyr	His	Ile	Ser	Leu	Ile	Asp	Ile	Gly	450	455	460	
Leu	Val	Leu	Glu	Tyr	Leu	Met	Gly	Gly	Ala	Tyr	Arg	Cys	Asn	Tyr	Thr	465	470	475	480
Arg	Lys	Asn	Phe	Arg	Thr	Leu	Tyr	Asn	Asn	Leu	Phe	Gly	Pro	Lys	Arg	485	490	495	
Pro	Lys	Ala	Leu	Lys	Leu	Leu	Gly	Met	Glu	Asp	Asp	Glu	Pro	Pro	Ala	500	505	510	
Lys	Gly	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Glu	Glu	Glu	Ile	Asp	Ile	515	520	525	
Asp	Val	Asp	Asp	Pro	Ala	Val	Ser	Arg	Phe	Gln	Tyr	Pro	Phe	His	Glu	530	535	540	

-11-

Leu	Met	Val	Trp	Ala	Val	Leu	Met	Lys	Arg	Gln	Lys	Met	Ala	Val	Phe
545					550					555					560
Leu	Trp	Gln	Arg	Gly	Glu	Glu	Ser	Met	Ala	Lys	Ala	Leu	Val	Ala	Cys
				565					570						575
Lys	Leu	Tyr	Lys	Ala	Met	Ala	His	Glu	Ser	Ser	Glu	Ser	Asp	Leu	Val
			580					585					590		
Asp	Asp	Ile	Ser	Gln	Asp	Leu	Asp	Asn	Asn	Ser	Lys	Asp	Phe	Gly	Gln
		595					600					605			
Leu	Ala	Leu	Glu	Leu	Leu	Asp	Gln	Ser	Tyr	Lys	His	Asp	Glu	Gln	Ile
	610					615					620				
Ala	Met	Lys	Leu	Leu	Thr	Tyr	Glu	Leu	Lys	Asn	Trp	Ser	Asn	Ser	Thr
625					630					635					640
Cys	Leu	Lys	Leu	Ala	Val	Ala	Ala	Lys	His	Arg	Asp	Phe	Ile	Ala	His
				645					650						655
Thr	Cys	Ser	Gln	Met	Leu	Leu	Thr	Asp	Met	Trp	Met	Gly	Arg	Leu	Arg
			660					665					670		
Met	Arg	Lys	Asn	Pro	Gly	Leu	Lys	Val	Ile	Met	Gly	Ile	Leu	Leu	Pro
		675					680					685			
Pro	Thr	Ile	Leu	Phe	Leu	Glu	Phe	Arg	Thr	Tyr	Asp	Asp	Phe	Ser	Tyr
	690					695					700				
Gln	Thr	Ser	Lys	Glu	Asn	Glu	Asp	Gly	Lys	Glu	Lys	Glu	Glu	Glu	Asn
705					710					715					720
Thr	Asp	Ala	Asn	Ala	Asp	Ala	Gly	Ser	Arg	Lys	Gly	Asp	Glu	Glu	Asn
			725						730						735
Glu	His	Lys	Lys	Gln	Arg	Ser	Ile	Pro	Ile	Gly	Thr	Lys	Ile	Cys	Glu
			740					745					750		
Phe	Tyr	Asn	Ala	Pro	Ile	Val	Lys	Phe	Trp	Phe	Tyr	Thr	Ile	Ser	Tyr
		755					760					765			
Leu	Gly	Tyr	Leu	Leu	Leu	Phe	Asn	Tyr	Val	Ile	Leu	Val	Arg	Met	Asp
	770					775					780				
Gly	Trp	Pro	Ser	Leu	Gln	Glu	Trp	Ile	Val	Ile	Ser	Tyr	Ile	Val	Ser
785					790					795					800
Leu	Ala	Leu	Glu	Lys	Ile	Arg	Glu	Ile	Leu	Met	Ser	Glu	Pro	Gly	Lys
				805					810						815
Leu	Ser	Gln	Lys	Ile	Lys	Val	Trp	Leu	Gln	Glu	Tyr	Trp	Asn	Ile	Thr
			820					825					830		
Asp	Leu	Val	Ala	Ile	Ser	Thr	Phe	Met	Ile	Gly	Ala	Ile	Leu	Arg	Leu
		835					840					845			
Gln	Asn	Gln	Pro	Tyr	Met	Gly	Tyr	Gly	Arg	Val	Ile	Tyr	Cys	Val	Asp
	850					855					860				
Ile	Ile	Phe	Trp	Tyr	Ile	Arg	Val	Leu	Asp	Ile	Phe	Gly	Val	Asn	Lys
865					870					875					880
Tyr	Leu	Gly	Pro	Tyr	Val	Met	Met	Ile	Gly	Lys	Met	Met	Ile	Asp	Met
				885					890					895	
Leu	Tyr	Phe	Val	Val	Ile	Met	Leu	Val	Val	Leu	Met	Ser	Phe	Gly	Val
			900					905					910		
Ala	Arg	Gln	Ala	Ile	Leu	His	Pro	Glu	Glu	Lys	Pro	Ser	Trp	Lys	Leu
			915				920						925		
Ala	Arg	Asn	Ile	Phe	Tyr	Met	Pro	Tyr	Trp	Met	Ile	Tyr	Gly	Glu	Val
						935					940				
Phe	Ala	Asp	Gln	Ile	Asp	Leu	Tyr	Ala	Met	Glu	Ile	Asn	Pro	Pro	Cys
945					950					955					960
Gly	Glu	Asn	Leu	Tyr	Asp	Glu	Glu	Gly	Lys	Arg	Leu	Pro	Pro	Cys	Ile
				965					970						975
Pro	Gly	Ala	Trp	Leu	Thr	Pro	Ala	Leu	Met	Ala	Cys	Tyr	Leu	Leu	Val
			980					985					990		
Ala	Asn	Ile	Leu	Leu	Val	Asn	Leu	Leu	Ile	Ala	Val	Phe	Asn	Asn	Thr
		995					1000					1005			
Phe	Phe	Glu	Val	Lys	Ser	Ile	Ser	Asn	Gln	Val	Trp	Lys	Phe	Gln	Arg
	1010						1015					1020			

-12-

Tyr Gln Leu Ile Met Thr Phe His Asp Arg Pro Val Leu Pro Pro Pro
 1025 1030 1035 104
 Met Ile Ile Leu Ser His Ile Tyr Ile Ile Met Arg Leu Ser Gly
 1045 1050 1055
 Arg Cys Arg Lys Lys Arg Glu Gly Asp Gln Glu Glu Arg Asp Arg Gly
 1060 1065 1070
 Leu Lys Leu Phe Leu Ser Asp Glu Glu Leu Lys Arg Leu His Glu Phe
 1075 1080 1085
 Glu Glu Gln Cys Val Gln Glu His Phe Arg Glu Lys Glu Asp Glu Gln
 1090 1095 1100
 Gln Ser Ser Ser Asp Glu Arg Ile Arg Val Thr Ser Glu Arg Val Glu
 1105 1110 1115 112
 Asn Met Ser Met Arg Leu Glu Glu Ile Asn Glu Arg Glu Thr Phe Met
 1125 1130 1135
 Lys Thr Ser Leu Gln Thr Val Asp Leu Arg Leu Ala Gln Leu Glu Glu
 1140 1145 1150
 Leu Ser Asn Arg Met Val Asn Ala Leu Glu Asn Leu Ala Gly Ile Asp
 1155 1160 1165
 Arg Ser Asp Leu Ile Gln Ala Arg Ser Arg Ala Ser Ser Glu Cys Glu
 1170 1175 1180
 Ala Thr Tyr Leu Leu Arg Gln Ser Ser Ile Asn Ser Ala Asp Gly Tyr
 1185 1190 1195 120
 Ser Leu Tyr Arg Tyr His Phe Asn Gly Glu Glu Leu Leu Phe Glu Asp
 1205 1210 1215
 Thr Ser Leu Ser Thr Ser Pro Gly Thr Gly Val Arg Lys Lys Thr Cys
 1220 1225 1230
 Ser Phe Arg Ile Lys Glu Glu Lys Asp Val Lys Thr His Leu Val Pro
 1235 1240 1245
 Glu Cys Gln Asn Ser Leu His Leu Ser Leu Gly Thr Ser Thr Ser Ala
 1250 1255 1260
 Thr Pro Asp Gly Ser His Leu Ala Val Asp Asp Leu Lys Asn Ala Glu
 1265 1270 1275 128
 Glu Ser Lys Leu Gly Pro Asp Ile Gly Ile Ser Lys Glu Asp Asp Glu
 1285 1290 1295
 Arg Gln Thr Asp Ser Lys Lys Glu Glu Thr Ile Ser Pro Ser Leu Asn
 1300 1305 1310
 Lys Thr Asp Val Ile His Gly Gln Asp Lys Ser Asp Val Gln Asn Thr
 1315 1320 1325
 Gln Leu Thr Val Glu Thr Thr Asn Ile Glu Gly Thr Ile Ser Tyr Pro
 1330 1335 1340
 Leu Glu Glu Thr Lys Ile Thr Arg Tyr Phe Pro Asp Glu Thr Ile Asn
 1345 1350 1355 136
 Ala Cys Lys Thr Met Lys Ser Arg Ser Phe Val Tyr Ser Arg Gly Arg
 1365 1370 1375
 Lys Leu Val Gly Gly Val Asn Gln Asp Val Glu Tyr Ser Ser Ile Thr
 1380 1385 1390
 Asp Gln Gln Leu Thr Thr Glu Trp Gln Cys Gln Val Gln Lys Ile Thr
 1395 1400 1405
 Arg Ser His Ser Thr Asp Ile Pro Tyr Ile Val Ser Glu Ala Ala Val
 1410 1415 1420
 Gln Ala Glu Gln Lys Glu Gln Phe Ala Asp Met Gln Asp Glu His His
 1425 1430 1435 144
 Val Ala Glu Ala Ile Pro Arg Ile Pro Arg Leu Ser Leu Thr Ile Thr
 1445 1450 1455
 Asp Arg Asn Gly Met Glu Asn Leu Leu Ser Val Lys Pro Asp Gln Thr
 1460 1465 1470
 Leu Gly Phe Pro Ser Leu Arg Ser Lys Ser Leu His Gly His Pro Arg
 1475 1480 1485
 Asn Val Lys Ser Ile Gln Gly Lys Leu Asp Arg Ser Gly His Ala Ser
 1490 1495 1500

-13-

Ser Val Ser Ser Leu Val Ile Val Ser Gly Met Thr Ala Glu Glu Lys
 1505 1510 1515 152
 Lys Val Lys Lys Glu Lys Ala Ser Thr Glu Thr Glu Cys
 1525 1530

<210> 11
 <211> 6220
 <212> DNA
 <213> Homo Sapiens

<400> 11
 tgtgcagaat tgtacagttg cgaaaccatg tcgctggcag ctggtgctgg cgggtggagac 60
 ttccctgtgc ggtgctcagt gcatctgcac ccgtggggga gggagctctt tctctggccc 120
 tgcagtcacc tgaggttggt accattatga acggccgctg ggacccccgc atgtgcatgt 180
 actccccag agtgtccggg ggccccagcc aagggaacaca tctcacgcag ctgggaacat 240
 gtgcaggctg atgaagagaa ccggatgagg gcttcacatg aggaagcatg tggccaggtc 300
 ctctcagaac atcagcctca tcttctgtc tctgatctat ttcaccaacc accccatgtg 360
 tctctagaac ccagtgtag cgagctggag agaggactgt cctgagggca gcaggcctgg 420
 ttgcagctgg cgtgggggtc tcagaatgga gccctcagcc ctgaggaaag ctggctcggg 480
 gcaggaggag ggctttgagg ggctgccag aagggtcact gacctgggga tgggtctccaa 540
 tctccggcgc agcaacagca gcctcttcaa gagctggagg ctacagtgcc ccttcggcaa 600
 caatgacaag caagaaagcc tcagttcgtg gattcctgaa aacatcaaga agaaagaatg 660
 cgtgtatttt gtggaaagt ccaaaactgt tgatgctggg aagggtgtgt gtcagtgtgg 720
 ctacacgcat gagcagcact tggaggaggc taccaagccc cacacctcc agggcacaca 780
 gtgggaccca aagaaacatg tccaggagat gccaaaccgat gcctttggcg acatcgtctt 840
 cacgggcctg agccagaagg tgaaaaagta cgtccgagtc tcccaggaca cgccctccag 900
 cgtgatctac caccctcatg ccagcactg ggggctggac gtccccaatc tcttgatctc 960
 ggtgaccggg ggggccaaga acttcaacat gaagccgcgg ctgaagagca ttttccgcag 1020
 aggcctggtc aagggtggctc agaccacagg ggcctggatc atcacagggg ggtcccacac 1080
 cggcgtcatg aagcaggtag gcgaggcggg gcgggacttc agcctgagca gcagctacaa 1140
 ggaaggcgag ctcatcacca tcggagtcgc cactctgggc actgtccacc gccgcgagg 1200
 cctgatccat ccacagggca gcttccccgc cgagtagata ctggatgagg atggccaagg 1260
 gaacctgacc tgcttagaca gcaaccactc tcacttcatc ctctgtagcg acgggaccca 1320
 cggccagtac ggggtggaga ttctctgag gaccaggctg gagaagttca tatcggagca 1380
 gaccaaggaa agaggagggt tggccatcaa gatccccatc gtgtgctggt tgggtggagg 1440
 cggcccgggc actgtgcaca ccacgcacaa gccaccacc aacggcacc cctgtgtggt 1500
 tgtggagggg tcgggcccgc tggccgacgt cattgccag gtggccaacc tgcctgtctc 1560
 ggacatcact atctccctga tccagcagaa actgagcgtg ttcttccagg agatgtttga 1620
 gaccttcacg gaaagcagga ttgtcgagt gacaaaaag atccaagata ttgtccggag 1680
 gcggcagctg ctagctgtct tccgggaagg caaggatgt cagcaggacg tggatgtggc 1740
 catcttgtag gccttgctga aagcctcacg gagccaagac cactttggcc acgagaactg 1800
 ggaccaccag ctgaaactgg cagtggcatg gaatcgctg gacattgccc gcagtggagt 1860
 ctctatggat gagtggcagt ggaagccttc agatctgcac cccacgatga cagctgact 1920
 catctccaac aagcctgagt ttgtgaagct ctctctgaa aacggggtgc agctgaagg 1980
 gtttgtcacc tgggacacct tgcctacct gtacgagaa ctggacccct cctgcctgtt 2040
 ccacagcaag ctgcaaaagg tgcgtggtga ggatcccag cggccggctt gcgcgcccgc 2100
 ggcgccccgc ctgcagatgc accacgtggc ccagggtgtg cgggagctgc tgggggactt 2160
 caccgagcgc ctttatcccc ggccccgca caacgaccg ctgcggtcc tgctgcccgt 2220
 tccccagctc aagctcaacg tgcaggagt gagcctccg tccctctaca agcgttctc 2280
 aggcctatg accttcacca tggaccccat ccgtgacct ctcatgtgg ccattgtcca 2340
 gaaccgtcg gagctggcag gaatcatctg ggctcagagc caggactgca tcgcagcggc 2400
 cttggcctgc agcaagatcc tgaaggaaat gtccaaggag gaggaggaca cggacagctc 2460
 ggaggagatg ctggcgctgg cggaggagta tgagcacaga gccatcgggg tcttcaccga 2520
 gtgctaccgg aaggacgaag agagagccca gaaactgtc acccgctgt cggaggcctg 2580
 ggggaagacc acctgcctgc agctcgccct ggaggccaag gacatgaaat ttgtgtctca 2640
 cgggggcatc caggccttcc tgaccaaggt gtggtggggc cagctctccg tggacaatgg 2700
 gctgtggcgt gtgacctgt gcatgctggc ctcccgctg ctctcaccg gcctcatctc 2760
 cttcagggag aagaggctgc aggatgtggg ccccccgcg gcccgcgccc gtgccttctt 2820
 caccgcaccc gtggtggtct tccacctgaa catctctcc tacttgcct tctctgcct 2880
 gttgcgctac gtgctcatgg tggacttcca gcctgtgcc tctggtgctg agtgtgccat 2940

-14-

ctacctctgg	ctcttctcct	tgggtgtgca	ggagatgcgg	cagctcttct	atgaccctga	3000
cgagtgcggg	ctgatgaaga	aggcagcctt	gtacttcagt	gacttctgga	ataagctgga	3060
cgctggcgca	atcttgctct	tcgtggcagg	gctgacctgc	aggctcatcc	cggcgacgct	3120
gtaccccggg	cgcgatcatc	tctctctgga	cttcatcctg	ttctgcctcc	ggctcatgca	3180
cattttttacc	atcagtaaga	cgctggggcc	caagatcatc	attgtgaagc	ggatgatgaa	3240
ggacgtcttc	ttcttctcct	tcctgctggc	tgtgtgggtg	gtgtccttcg	gggtggccaa	3300
gcaggccatc	ctcatccaca	acgagcgccg	ggtggactgg	ctgttccgag	gggcccgtct	3360
ccactcctac	ctcaccatct	tcgggcagat	cccgggttac	atcgacggtg	tgaacttcaa	3420
cccggagcac	tgcagcccca	atggcaccga	cccctacaag	cctaagtgcc	ccgagagcga	3480
cgcgacgcag	cagaggccgg	ccttccctga	gtggctgacg	gtcctcctac	tctgcctcta	3540
cctgctcttc	accaacatcc	tgctgctcaa	cctcctcatc	gccatgttca	actacacctt	3600
ccagcaggtg	caggagcaca	cggaccagat	ttggaagttc	cagcgccatg	acctgatcga	3660
ggagtaccac	ggcgccccc	cccgccccc	cccttcatc	ctcctcagcc	acctgcagct	3720
cttcatcaag	agggtggtcc	tgaagactcc	ggccaagagg	cacaagcagc	tcaagaacaa	3780
gctggagaag	aacgaggagg	cggccctgct	atcctgggag	atctacctga	aggagaacta	3840
cctccagaac	cgacagtcc	agcaaaagca	gcgcccgag	cagaagatcg	aggacatcag	3900
caataaggtt	gacgccatgg	tggacctgct	ggacctggac	ccactgaaga	ggtcgggctc	3960
catggagcag	agggtggcct	ccctggagga	gcaggtggcc	cagacagccc	gagccctgca	4020
ctggatcgtg	aggacgctgc	gggcccagcg	cttcagctcg	gaggcggacg	tccccactct	4080
ggcctcccag	aaggccgagg	aggagccgga	tgctgagccg	ggaggcagga	agaagacgga	4140
ggagccgggc	gacagctacc	acgtgaatgc	ccggcacctc	ctctacccca	actgccctgt	4200
cacgcgcttc	cccggtccca	acgagaaggt	gccttgggag	acggagtctc	tgatctatga	4260
cccacccttt	tacacggcag	agaggaagga	cgcggccgcc	atggacccca	tgggagacac	4320
cctggagcca	ctgtccacga	tccagtacaa	cgtggtggat	ggcctgaggg	accgccggag	4380
cttccacggg	ccgtacacag	tgcaggccgg	gttgcccctg	aaccccatgg	gccgcacagg	4440
actgcgtggg	cgcgggagcc	tcagctgctt	cggaccacaac	cacacgctgt	accccatggg	4500
cacgcggtgg	aggcggaacg	aggatggagc	catctgcagg	aagagcataa	agaagatgct	4560
ggaagtgcctg	gtggtgaagc	tccctctctc	cgagcactgg	gccctgcctg	ggggctcccg	4620
ggagccaggg	gagatgtctac	ctcggaagct	gaagcggatc	ctccggcagg	agcactggcc	4680
gtcttttgaa	aacttgctga	agtgcggcat	ggaggtgtac	aaaggctaca	tggatgaccc	4740
gaggaacacg	gacaatgcct	ggatcgagac	ggtggccgtc	agcgtccact	tccaggacca	4800
gaatgacgtg	gagctgaaca	ggctgaactc	taacctgcac	gcctgcgact	cgggggccc	4860
catccgatgg	cagggtggtg	acaggcgcat	cccactctat	gcgaaccaca	agaccctcct	4920
ccagaaggca	gcccgtagt	tcggggctca	ctactgactg	tgccctcagg	ctgggcggct	4980
ccagtccata	cagcttcccc	ccagaaacca	gtcttctct	ctcctgagcc	tggccaggac	5040
tcaggctggt	cctgggcccct	gcacatgatg	gggtttggtg	gacccagtgc	ccctcacggc	5100
tgccgcaagt	ctgctgcaga	tgacctcatg	aactggaagg	ggtcaagggtg	acccgggagg	5160
agagctcaag	acagggcaca	ggctactcag	agctgagggg	cccctgggac	ccttggccat	5220
caggcgaggg	gctgggcctg	tgcagctggg	cccttggcca	gagtccactc	ccttctctggc	5280
tgtgtcaccc	cgagcagctc	atccaccatg	gaggtcattg	gcctgaggca	agttccccgg	5340
agagtgggga	tcccctgtgg	ccccctcagg	cctatgtctg	tgagggaagg	gccctgccac	5400
tctccccaag	agggcctcca	tgtttcgagg	tgccctcaaca	tggagccttg	cctggccttg	5460
gctaggggca	ctgtctgaac	tcctgactgt	caggataaac	tccgtggggg	tacaggagcc	5520
cagacaaagc	ccaggcctgt	caagagacgc	agagggcccc	tgccagggtt	ggccccaggg	5580
accctgggac	gaggctgcag	aagctctccc	tccctactcc	ctgggagcca	cgtgctggcc	5640
atgtggccag	ggacggcatg	agcaggaggc	ggggacgtgg	gggccttctg	gtttggtgtc	5700
aacagctcac	aggagcgtga	accatgaggg	ccctcaggag	gggaacgtgg	taaaacccaa	5760
gacattaaat	ctgccatctc	aggcctggct	ggctcttctg	tgctttccac	aaataaagtt	5820
cctgacacgt	ccaggggccag	gggctgtgtg	acggctgcct	gaagtctctc	tcgatccccc	5880
ggtgagcttc	ctgcagcctg	tggatgtcct	gcagcccctc	agccctaccc	ccaagtctct	5940
cctctgaccc	atcagctccc	tgtcttcatt	ttcctaaacc	tgggctccag	catcgtcccc	6000
aagcccacca	ggccgggatg	caggcatcca	catgcctcc	tccttggctt	cccctgcgtg	6060
gtgggtgccaa	tgtgccctgg	cacccctgca	gaggctccgg	atggagcctg	gggctgcctg	6120
gccactgagc	actggccgag	gtgatgccca	cccttcctg	gacaggcctc	tgtcttccac	6180
ctgacccaaa	gctctctagc	caccccttg	tccccagtat			6220

<210> 12

<211> 1503

-15-

<212> PRT

<213> Homo Sapiens

<400> 12

Met	Glu	Pro	Ser	Ala	Leu	Arg	Lys	Ala	Gly	Ser	Glu	Gln	Glu	Glu	Gly
1				5					10				15		
Phe	Glu	Gly	Leu	Pro	Arg	Arg	Val	Thr	Asp	Leu	Gly	Met	Val	Ser	Asn
			20					25					30		
Leu	Arg	Arg	Ser	Asn	Ser	Ser	Leu	Phe	Lys	Ser	Trp	Arg	Leu	Gln	Cys
			35				40					45			
Pro	Phe	Gly	Asn	Asn	Asp	Lys	Gln	Glu	Ser	Leu	Ser	Ser	Trp	Ile	Pro
			50			55					60				
Glu	Asn	Ile	Lys	Lys	Lys	Glu	Cys	Val	Tyr	Phe	Val	Glu	Ser	Ser	Lys
65						70				75					80
Leu	Ser	Asp	Ala	Gly	Lys	Val	Val	Cys	Gln	Cys	Gly	Tyr	Thr	His	Glu
				85					90					95	
Gln	His	Leu	Glu	Ala	Thr	Lys	Pro	His	Thr	Phe	Gln	Gly	Thr	Gln	
			100				105					110			
Trp	Asp	Pro	Lys	Lys	His	Val	Gln	Glu	Met	Pro	Thr	Asp	Ala	Phe	Gly
			115				120					125			
Asp	Ile	Val	Phe	Thr	Gly	Leu	Ser	Gln	Lys	Val	Lys	Lys	Tyr	Val	Arg
			130			135					140				
Val	Ser	Gln	Asp	Thr	Pro	Ser	Ser	Val	Ile	Tyr	His	Leu	Met	Thr	Gln
145					150					155					160
His	Trp	Gly	Leu	Asp	Val	Pro	Asn	Leu	Leu	Ile	Ser	Val	Thr	Gly	Gly
				165					170					175	
Ala	Lys	Asn	Phe	Asn	Met	Lys	Pro	Arg	Leu	Lys	Ser	Ile	Phe	Arg	Arg
			180					185					190		
Gly	Leu	Val	Lys	Val	Ala	Gln	Thr	Thr	Gly	Ala	Trp	Ile	Ile	Thr	Gly
			195				200					205			
Gly	Ser	His	Thr	Gly	Val	Met	Lys	Gln	Val	Gly	Glu	Ala	Val	Arg	Asp
			210			215					220				
Phe	Ser	Leu	Ser	Ser	Ser	Tyr	Lys	Glu	Gly	Glu	Leu	Ile	Thr	Ile	Gly
225					230					235					240
Val	Ala	Thr	Trp	Gly	Thr	Val	His	Arg	Arg	Glu	Gly	Leu	Ile	His	Pro
				245					250					255	
Thr	Gly	Ser	Phe	Pro	Ala	Glu	Tyr	Ile	Leu	Asp	Glu	Asp	Gly	Gln	Gly
			260					265					270		
Asn	Leu	Thr	Cys	Leu	Asp	Ser	Asn	His	Ser	His	Phe	Ile	Leu	Val	Asp
			275				280					285			
Asp	Gly	Thr	His	Gly	Gln	Tyr	Gly	Val	Glu	Ile	Pro	Leu	Arg	Thr	Arg
			290			295					300				
Leu	Glu	Lys	Phe	Ile	Ser	Glu	Gln	Thr	Lys	Glu	Arg	Gly	Gly	Val	Ala
305					310					315					320
Ile	Lys	Ile	Pro	Ile	Val	Cys	Val	Val	Leu	Glu	Gly	Gly	Pro	Gly	Thr
				325					330					335	
Leu	His	Thr	Ile	Asp	Asn	Ala	Thr	Thr	Asn	Gly	Thr	Pro	Cys	Val	Val
			340					345					350		
Val	Glu	Gly	Ser	Gly	Arg	Val	Ala	Asp	Val	Ile	Ala	Gln	Val	Ala	Asn
			355				360					365			
Leu	Pro	Val	Ser	Asp	Ile	Thr	Ile	Ser	Leu	Ile	Gln	Gln	Lys	Leu	Ser
			370			375					380				
Val	Phe	Phe	Gln	Glu	Met	Phe	Glu	Thr	Phe	Thr	Glu	Ser	Arg	Ile	Val
385					390					395					400
Glu	Trp	Thr	Lys	Lys	Ile	Gln	Asp	Ile	Val	Arg	Arg	Arg	Gln	Leu	Leu
				405					410					415	
Thr	Val	Phe	Arg	Glu	Gly	Lys	Asp	Gly	Gln	Gln	Asp	Val	Asp	Val	Ala
			420					425					430		
Ile	Leu	Gln	Ala	Leu	Leu	Lys	Ala	Ser	Arg	Ser	Gln	Asp	His	Phe	Gly
			435				440					445			

-16-

His	Glu	Asn	Trp	Asp	His	Gln	Leu	Lys	Leu	Ala	Val	Ala	Trp	Asn	Arg
450					455					460					
Val	Asp	Ile	Ala	Arg	Ser	Glu	Ile	Phe	Met	Asp	Glu	Trp	Gln	Trp	Lys
465					470					475					480
Pro	Ser	Asp	Leu	His	Pro	Thr	Met	Thr	Ala	Ala	Leu	Ile	Ser	Asn	Lys
				485					490						495
Pro	Glu	Phe	Val	Lys	Leu	Phe	Leu	Glu	Asn	Gly	Val	Gln	Leu	Lys	Glu
			500					505					510		
Phe	Val	Thr	Trp	Asp	Thr	Leu	Leu	Tyr	Leu	Tyr	Glu	Asn	Leu	Asp	Pro
		515				520						525			
Ser	Cys	Leu	Phe	His	Ser	Lys	Leu	Gln	Lys	Val	Leu	Val	Glu	Asp	Pro
530					535						540				
Glu	Arg	Pro	Ala	Cys	Ala	Pro	Ala	Ala	Pro	Arg	Leu	Gln	Met	His	His
545				550						555					560
Val	Ala	Gln	Val	Leu	Arg	Glu	Leu	Leu	Gly	Asp	Phe	Thr	Gln	Pro	Leu
				565					570						575
Tyr	Pro	Arg	Pro	Arg	His	Asn	Asp	Arg	Leu	Arg	Leu	Leu	Leu	Pro	Val
			580					585					590		
Pro	His	Val	Lys	Leu	Asn	Val	Gln	Gly	Val	Ser	Leu	Arg	Ser	Leu	Tyr
		595				600						605			
Lys	Arg	Ser	Ser	Gly	His	Val	Thr	Phe	Thr	Met	Asp	Pro	Ile	Arg	Asp
610					615						620				
Leu	Leu	Ile	Trp	Ala	Ile	Val	Gln	Asn	Arg	Arg	Glu	Leu	Ala	Gly	Ile
625				630						635					640
Ile	Trp	Ala	Gln	Ser	Gln	Asp	Cys	Ile	Ala	Ala	Ala	Leu	Ala	Cys	Ser
				645					650					655	
Lys	Ile	Leu	Lys	Glu	Leu	Ser	Lys	Glu	Glu	Glu	Asp	Thr	Asp	Ser	Ser
			660					665					670		
Glu	Glu	Met	Leu	Ala	Leu	Ala	Glu	Glu	Tyr	Glu	His	Arg	Ala	Ile	Gly
		675					680					685			
Val	Phe	Thr	Glu	Cys	Tyr	Arg	Lys	Asp	Glu	Glu	Arg	Ala	Gln	Lys	Leu
		690				695					700				
Leu	Thr	Arg	Val	Ser	Glu	Ala	Trp	Gly	Lys	Thr	Thr	Cys	Leu	Gln	Leu
705				710						715					720
Ala	Leu	Glu	Ala	Lys	Asp	Met	Lys	Phe	Val	Ser	His	Gly	Gly	Ile	Gln
				725					730					735	
Ala	Phe	Leu	Thr	Lys	Val	Trp	Trp	Gly	Gln	Leu	Ser	Val	Asp	Asn	Gly
			740					745					750		
Leu	Trp	Arg	Val	Thr	Leu	Cys	Met	Leu	Ala	Phe	Pro	Leu	Leu	Leu	Thr
		755					760					765			
Gly	Leu	Ile	Ser	Phe	Arg	Glu	Lys	Arg	Leu	Gln	Asp	Val	Gly	Thr	Pro
770						775					780				
Ala	Ala	Arg	Ala	Arg	Ala	Phe	Phe	Thr	Ala	Pro	Val	Val	Val	Phe	His
785					790					795					800
Leu	Asn	Ile	Leu	Ser	Tyr	Phe	Ala	Phe	Leu	Cys	Leu	Phe	Ala	Tyr	Val
				805					810					815	
Leu	Met	Val	Asp	Phe	Gln	Pro	Val	Pro	Ser	Trp	Cys	Glu	Cys	Ala	Ile
			820					825					830		
Tyr	Leu	Trp	Leu	Phe	Ser	Leu	Val	Cys	Glu	Glu	Met	Arg	Gln	Leu	Phe
		835					840					845			
Tyr	Asp	Pro	Asp	Glu	Cys	Gly	Leu	Met	Lys	Lys	Ala	Ala	Leu	Tyr	Phe
		850				855					860				
Ser	Asp	Phe	Trp	Asn	Lys	Leu	Asp	Val	Gly	Ala	Ile	Leu	Leu	Phe	Val
865				870						875					880
Ala	Gly	Leu	Thr	Cys	Arg	Leu	Ile	Pro	Ala	Thr	Leu	Tyr	Pro	Gly	Arg
				885					890					895	
Val	Ile	Leu	Ser	Leu	Asp	Phe	Ile	Leu	Phe	Cys	Leu	Arg	Leu	Met	His
			900					905					910		
Ile	Phe	Thr	Ile	Ser	Lys	Thr	Leu	Gly	Pro	Lys	Ile	Ile	Ile	Val	Lys
		915					920						925		

-17-

Arg	Met	Met	Lys	Asp	Val	Phe	Phe	Phe	Leu	Phe	Leu	Leu	Ala	Val	Trp
	930					935					940				
Val	Val	Ser	Phe	Gly	Val	Ala	Lys	Gln	Ala	Ile	Leu	Ile	His	Asn	Glu
945					950					955					960
Arg	Arg	Val	Asp	Trp	Leu	Phe	Arg	Gly	Ala	Val	Tyr	His	Ser	Tyr	Leu
				965					970					975	
Thr	Ile	Phe	Gly	Gln	Ile	Pro	Gly	Tyr	Ile	Asp	Gly	Val	Asn	Phe	Asn
			980					985					990		
Pro	Glu	His	Cys	Ser	Pro	Asn	Gly	Thr	Asp	Pro	Tyr	Lys	Pro	Lys	Cys
		995					1000					1005			
Pro	Glu	Ser	Asp	Ala	Thr	Gln	Gln	Arg	Pro	Ala	Phe	Pro	Glu	Trp	Leu
	1010					1015				1020					
Thr	Val	Leu	Leu	Leu	Cys	Leu	Tyr	Leu	Leu	Phe	Thr	Asn	Ile	Leu	Leu
1025					1030					1035					104
Leu	Asn	Leu	Leu	Ile	Ala	Met	Phe	Asn	Tyr	Thr	Phe	Gln	Gln	Val	Gln
				1045					1050					1055	
Glu	His	Thr	Asp	Gln	Ile	Trp	Lys	Phe	Gln	Arg	His	Asp	Leu	Ile	Glu
		1060						1065				1070			
Glu	Tyr	His	Gly	Arg	Pro	Ala	Ala	Pro	Pro	Pro	Phe	Ile	Leu	Leu	Ser
	1075					1080					1085				
His	Leu	Gln	Leu	Phe	Ile	Lys	Arg	Val	Val	Leu	Lys	Thr	Pro	Ala	Lys
	1090					1095					1100				
Arg	His	Lys	Gln	Leu	Lys	Asn	Lys	Leu	Glu	Lys	Asn	Glu	Glu	Ala	Ala
1105					1110					1115					112
Leu	Leu	Ser	Trp	Glu	Ile	Tyr	Leu	Lys	Glu	Asn	Tyr	Leu	Gln	Asn	Arg
				1125					1130					1135	
Gln	Phe	Gln	Gln	Lys	Gln	Arg	Pro	Glu	Gln	Lys	Ile	Glu	Asp	Ile	Ser
				1140				1145				1150			
Asn	Lys	Val	Asp	Ala	Met	Val	Asp	Leu	Leu	Asp	Leu	Asp	Pro	Leu	Lys
	1155						1160				1165				
Arg	Ser	Gly	Ser	Met	Glu	Gln	Arg	Leu	Ala	Ser	Leu	Glu	Glu	Gln	Val
	1170				1175					1180					
Ala	Gln	Thr	Ala	Arg	Ala	Leu	His	Trp	Ile	Val	Arg	Thr	Leu	Arg	Ala
1185					1190					1195					120
Ser	Gly	Phe	Ser	Ser	Glu	Ala	Asp	Val	Pro	Thr	Leu	Ala	Ser	Gln	Lys
				1205					1210					1215	
Ala	Ala	Glu	Glu	Pro	Asp	Ala	Glu	Pro	Gly	Gly	Arg	Lys	Lys	Thr	Glu
	1220						1225					1230			
Glu	Pro	Gly	Asp	Ser	Tyr	His	Val	Asn	Ala	Arg	His	Leu	Leu	Tyr	Pro
	1235						1240				1245				
Asn	Cys	Pro	Val	Thr	Arg	Phe	Pro	Val	Pro	Asn	Glu	Lys	Val	Pro	Trp
	1250					1255				1260					
Glu	Thr	Glu	Phe	Leu	Ile	Tyr	Asp	Pro	Pro	Phe	Tyr	Thr	Ala	Glu	Arg
1265					1270					1275					128
Lys	Asp	Ala	Ala	Ala	Met	Asp	Pro	Met	Gly	Asp	Thr	Leu	Glu	Pro	Leu
				1285					1290					1295	
Ser	Thr	Ile	Gln	Tyr	Asn	Val	Val	Asp	Gly	Leu	Arg	Asp	Arg	Arg	Ser
		1300					1305					1310			
Phe	His	Gly	Pro	Tyr	Thr	Val	Gln	Ala	Gly	Leu	Pro	Leu	Asn	Pro	Met
	1315						1320					1325			
Gly	Arg	Thr	Gly	Leu	Arg	Gly	Arg	Gly	Ser	Leu	Ser	Cys	Phe	Gly	Pro
	1330				1335					1340					
Asn	His	Thr	Leu	Tyr	Pro	Met	Val	Thr	Arg	Trp	Arg	Arg	Asn	Glu	Asp
1345					1350					1355					136
Gly	Ala	Ile	Cys	Arg	Lys	Ser	Ile	Lys	Lys	Met	Leu	Glu	Val	Leu	Val
				1365					1370					1375	
Val	Lys	Leu	Pro	Leu	Ser	Glu	His	Trp	Ala	Leu	Pro	Gly	Gly	Ser	Arg
		1380					1385					1390			
Glu	Pro	Gly	Glu	Met	Leu	Pro	Arg	Lys	Leu	Lys	Arg	Ile	Leu	Arg	Gln
	1395						1400					1405			

-18-

Glu His Trp Pro Ser Phe Glu Asn Leu Leu Lys Cys Gly Met Glu Val
 1410 1415 1420
 Tyr Lys Gly Tyr Met Asp Asp Pro Arg Asn Thr Asp Asn Ala Trp Ile
 1425 1430 1435 144
 Glu Thr Val Ala Val Ser Val His Phe Gln Asp Gln Asn Asp Val Glu
 1445 1450 1455
 Leu Asn Arg Leu Asn Ser Asn Leu His Ala Cys Asp Ser Gly Ala Ser
 1460 1465 1470
 Ile Arg Trp Gln Val Val Asp Arg Arg Ile Pro Leu Tyr Ala Asn His
 1475 1480 1485
 Lys Thr Leu Leu Gln Lys Ala Ala Ala Glu Phe Gly Ala His Tyr
 1490 1495 1500

<210> 13
 <211> 1816
 <212> PRT
 <213> C. Elegans

<400> 13

Met Ile Thr Asp Lys Asn Leu Phe Ser Arg Leu Leu Ile Lys Lys Asn
 1 5 10 15
 Pro Ile Arg Met His Ser Pro Ser Phe Ser Phe Ser Leu Ile Thr Ser
 20 25 30
 Leu Phe Phe Thr Gln Phe Phe Met Phe Gln Leu Ser Ser Met Ala Tyr
 35 40 45
 Phe Phe Leu Thr Leu Ile Ala Gly Val Thr His Phe Tyr Phe Pro Glu
 50 55 60
 Lys Leu Leu Gly Lys Ser Glu Asn Leu Asp His Arg Tyr Gln Ser Ser
 65 70 75 80
 Glu Gln Lys Val Leu Ile Glu Trp Thr Glu Asn Lys Ala Val Ala Glu
 85 90 95
 Ser Leu Arg Ala Asn Ser Val Thr Val Glu Glu Asn Glu Ser Glu Arg
 100 105 110
 Glu Thr Glu Thr Gln Thr Lys Arg Arg Arg Lys Lys Gln Arg Ser Thr
 115 120 125
 Ser Ser Asp Lys Ala Pro Leu Asn Ser Ala Pro Arg His Val Gln Lys
 130 135 140
 Phe Asp Trp Lys Asp Met Leu His Leu Ala Asp Ile Ser Gly Arg Lys
 145 150 155 160
 Arg Gly Asn Ser Thr Ser His Ser Gly His Ala Thr Arg Ala Gly
 165 170 175
 Ser Leu Lys Gly Lys Asn Trp Ile Glu Cys Arg Leu Lys Met Arg Gln
 180 185 190
 Cys Ser Tyr Phe Val Pro Ser Gln Arg Phe Ser Glu Arg Cys Gly Cys
 195 200 205
 Gly Lys Glu Arg Ser Lys His Thr Glu Glu Val Leu Glu Arg Ser Gln
 210 215 220
 Asn Lys Asn His Pro Leu Asn His Leu Thr Leu Pro Gly Ile His Glu
 225 230 235 240
 Val Asp Thr Thr Asp Ala Asp Ala Asp Asp Asn Glu Val Asn Leu Thr
 245 250 255
 Pro Gly Arg Trp Ser Ile Gln Ser His Thr Glu Ile Val Pro Thr Asp
 260 265 270
 Ala Tyr Gly Asn Ile Val Phe Glu Gly Thr Ala His His Ala Gln Tyr
 275 280 285
 Ala Arg Ile Ser Phe Asp Ser Asp Pro Arg Asp Ile Val His Leu Met
 290 295 300
 Met Lys Val Trp Lys Leu Lys Pro Pro Lys Leu Ile Ile Thr Ile Asn
 305 310 315 320
 Gly Gly Leu Thr Lys Phe Asp Leu Gln Pro Lys Leu Ala Arg Thr Phe

				325					330					335	
Arg	Lys	Gly	Ile	Met	Lys	Ile	Ala	Lys	Ser	Thr	Asp	Ala	Trp	Ile	Ile
			340					345					350		
Thr	Ser	Gly	Leu	Asp	Glu	Gly	Val	Val	Lys	His	Leu	Asp	Ser	Ala	Leu
		355					360					365			
His	Ala	Leu	Glu	Phe	Trp	Ser	Phe	Gly	Leu	Phe	Trp	Val	Ile	Gln	Leu
	370					375					380				
Asp	Val	Leu	Leu	Ala	His	Ser	Met	Phe	Ile	Pro	Arg	Gly	Ser	Leu	Phe
385					390					395					400
Asp	His	Gly	Asn	His	Thr	Ser	Lys	Asn	His	Val	Val	Ala	Ile	Gly	Ile
			405						410					415	
Ala	Ser	Trp	Gly	Met	Leu	Lys	Gln	Arg	Ser	Arg	Phe	Val	Gly	Lys	Asp
			420					425					430		
Ser	Thr	Val	Thr	Tyr	Ala	Thr	Asn	Val	Phe	Asn	Asn	Thr	Arg	Leu	Lys
		435					440					445			
Glu	Leu	Asn	Asp	Asn	His	Ser	Tyr	Phe	Leu	Phe	Ser	Asp	Asn	Gly	Thr
		450					455					460			
Val	Asn	Arg	Tyr	Gly	Ala	Glu	Ile	Ile	Met	Arg	Lys	Arg	Leu	Glu	Ala
465					470					475					480
Tyr	Leu	Ala	Gln	Gly	Asp	Lys	Lys	Arg	Ser	Ala	Ile	Pro	Leu	Val	Cys
				485					490					495	
Val	Val	Leu	Glu	Gly	Gly	Ala	Phe	Thr	Ile	Lys	Met	Val	His	Asp	Tyr
			500					505					510		
Val	Thr	Thr	Ile	Pro	Arg	Ile	Pro	Val	Ile	Val	Cys	Asp	Gly	Ser	Gly
		515					520					525			
Arg	Ala	Ala	Asp	Ile	Leu	Ala	Phe	Ala	His	Gln	Ala	Val	Ser	Gln	Asn
		530				535					540				
Gly	Phe	Leu	Ser	Asp	Asn	Ile	Arg	Asn	Gln	Leu	Val	Asn	Ile	Val	Arg
545					550					555					560
Arg	Ile	Phe	Gly	Tyr	Asp	Pro	Lys	Thr	Ala	Gln	Lys	Leu	Ile	Lys	Gln
				565					570					575	
Ile	Val	Glu	Cys	Ser	Thr	Asn	Lys	Ser	Leu	Met	Thr	Ile	Phe	Arg	Leu
			580					585					590		
Gly	Glu	Ser	Ser	Arg	Glu	Asp	Leu	Asp	His	Val	Ile	Met	Ser	Cys	Leu
		595					600					605			
Leu	Lys	Gly	Gln	Asn	Leu	Ser	Pro	Pro	Glu	Gln	Leu	Gln	Leu	Ala	Leu
		610				615					620				
Ala	Trp	Asn	Arg	Ala	Asp	Ile	Ala	Arg	Thr	Glu	Ile	Phe	Ala	Asn	Gly
625					630					635					640
Thr	Glu	Trp	Thr	Thr	Gln	Asp	Leu	His	Asn	Ala	Met	Ile	Glu	Ala	Leu
				645					650					655	
Ser	Asn	Asp	Arg	Ile	Asp	Phe	Val	His	Leu	Leu	Leu	Glu	Asn	Gly	Val
			660					665					670		
Ser	Met	Gln	Lys	Phe	Leu	Thr	Tyr	Gly	Arg	Leu	Glu	His	Leu	Tyr	Asn

-20-

				805					810					815	
Asn	Met	Asp	Phe	Thr	Phe	Arg	Tyr	Pro	Tyr	Ser	Asp	Leu	Met	Ile	Trp
			820						825				830		
Ala	Val	Leu	Thr	Lys	Arg	Gln	Lys	Met	Ala	Lys	Leu	Met	Trp	Thr	His
		835						840				845			
Gly	Glu	Glu	Gly	Met	Ala	Lys	Ala	Leu	Val	Ala	Ser	Arg	Leu	Tyr	Val
		850				855					860				
Ser	Leu	Ala	Lys	Thr	Ala	Ser	Leu	Ala	Thr	Gly	Glu	Ile	Gly	Met	Ser
865					870					875					880
Gln	Asp	Phe	Thr	Glu	Phe	Ser	Asp	Glu	Phe	Ser	Glu	Leu	Ala	Val	Glu
			885						890					895	
Val	Leu	Glu	Tyr	Cys	Thr	Lys	His	Gly	Arg	Asp	Gln	Thr	Leu	Arg	Leu
		900						905					910		
Leu	Thr	Cys	Glu	Leu	Ala	Asn	Trp	Gly	Asp	Glu	Thr	Cys	Leu	Ser	Leu
		915					920					925			
Ala	Ala	Asn	Asn	Gly	His	Arg	Lys	Phe	Leu	Ala	His	Pro	Cys	Cys	Gln
	930					935					940				
Met	Leu	Leu	Ser	Asp	Leu	Trp	Gln	Gly	Gly	Leu	Leu	Met	Lys	Asn	Asn
945					950					955					960
Gln	Asn	Ser	Lys	Val	Leu	Thr	Cys	Leu	Ala	Ala	Pro	Pro	Leu	Ile	Phe
			965						970					975	
Leu	Leu	Gly	Phe	Lys	Thr	Lys	Glu	Gln	Leu	Met	Leu	Gln	Pro	Lys	Thr
		980						985					990		
Ala	Ala	Glu	His	Asp	Glu	Glu	Met	Ser	Asp	Ser	Glu	Met	Asn	Ser	Ala
	995						1000					1005			
Glu	Asp	Thr	Asp	Thr	Ser	Ser	Asp	Ser	Ser	Ser	Asp	Ser	Asp	Asp	Ser
	1010					1015					1020				
Asp	Glu	Glu	Asp	Ala	Lys	Leu	Arg	Ala	Gln	Ser	Leu	Ser	Ala	Asp	Gln
1025					1030					1035					104
Pro	Leu	Ser	Ile	His	Arg	Leu	Val	Arg	Asp	Lys	Leu	Asn	Phe	Ser	Glu
			1045						1050					1055	
Lys	Lys	Lys	Pro	Asp	Met	Gly	Ile	Ser	Arg	Ile	Val	Val	Ala	Pro	Pro
			1060						1065				1070		
Ile	Val	Thr	Gly	Arg	Asn	Arg	Ala	Arg	Thr	Met	Ser	Ile	Lys	Lys	Ser
	1075						1080					1085			
Lys	Lys	Asn	Val	Ile	Lys	Pro	Pro	Ala	Cys	Leu	Lys	Ile	Glu	Thr	Ser
	1090					1095					1100				
Asp	Asp	Asp	Glu	Gln	Glu	Gln	Lys	Lys	Ala	Thr	Glu	Met	Cys	Lys	Ser
1105					1110					1115					112
Thr	Phe	Phe	Asp	Phe	Phe	Phe	Asp	Phe	Pro	Tyr	Ile	Asn	Arg	Thr	Gly
			1125						1130					1135	
Lys	Arg	Gly	Ser	Val	Ala	Val	Ala	Met	Asn	His	Asp	Asp	Met	Tyr	Ile
		1140						1145					1150		
Asp	Pro	Ser	Glu	Glu	Leu	Asp	Thr	Gln	Thr	Arg	Gln	Lys	Ser	Ser	Arg
		1155					1160					1165			
Glu	Phe	Ser	Ser	Ser	Arg	Asn	Val	Thr	Val	Gln	Val	Tyr	Thr	Gln	Arg
	1170					1175					1180				
Pro	Leu	Ser	Trp	Lys	Lys	Lys	Ile	Met	Glu	Phe	Tyr	Lys	Ala	Pro	Ile
1185					1190					1195					120
Thr	Thr	Tyr	Trp	Leu	Trp	Phe	Phe	Ala	Phe	Ile	Trp	Phe	Leu	Ile	Leu
			1205						1210					1215	
Leu	Thr	Tyr	Asn	Leu	Leu	Val	Lys	Thr	Gln	Arg	Ile	Ala	Ser	Trp	Ser
		1220						1225					1230		
Glu	Trp	Tyr	Val	Phe	Ala	Tyr	Ile	Phe	Val	Trp	Thr	Leu	Glu	Ile	Gly
		1235					1240					1245			
Arg	Lys	Val	Val	Ser	Thr	Ile	Met	Met	Asp	Thr	Ser	Lys	Pro	Val	Leu
	1250					1255					1260				
Lys	Gln	Leu	Arg	Val	Phe	Phe	Phe	Gln	Tyr	Arg	Asn	Gly	Leu	Leu	Ala
1265					1270					1275					128
Phe	Gly	Leu	Leu	Thr	Tyr	Leu	Ile	Ala	Tyr	Phe	Ile	Arg	Leu	Ser	Pro

-21-

				1285					1290					1295	
Thr	Thr	Lys	Thr	Leu	Gly	Arg	Ile	Leu	Ile	Ile	Cys	Asn	Ser	Val	Ile
				1300					1305					1310	
Trp	Ser	Leu	Lys	Leu	Val	Asp	Tyr	Leu	Ser	Val	Gln	Gln	Gly	Leu	Gly
				1315					1320					1325	
Pro	Tyr	Ile	Asn	Ile	Val	Ala	Glu	Met	Ile	Pro	Thr	Met	Ile	Pro	Leu
				1330					1335					1340	
Cys	Val	Leu	Val	Phe	Ile	Thr	Leu	Tyr	Ala	Phe	Gly	Leu	Leu	Arg	Gln
				1345					1350					1355	
Ser	Ile	Thr	Tyr	Pro	Tyr	Glu	Asp	Trp	His	Trp	Ile	Leu	Val	Arg	Asn
				1365					1370					1375	
Ile	Phe	Leu	Gln	Pro	Tyr	Phe	Met	Leu	Tyr	Gly	Glu	Val	Tyr	Ala	Ala
				1380					1385					1390	
Glu	Ile	Asp	Thr	Cys	Gly	Asp	Glu	Ile	Trp	Gln	Thr	His	Glu	Asp	Glu
				1395					1400					1405	
Asn	Ile	Pro	Ile	Ser	Met	Leu	Asn	Val	Thr	His	Glu	Thr	Cys	Val	Pro
				1410					1415					1420	
Gly	Tyr	Trp	Ile	Ala	Pro	Val	Gly	Leu	Thr	Val	Phe	Met	Leu	Ala	Thr
				1425					1430					1435	
Asn	Val	Leu	Leu	Met	Asn	Val	Met	Val	Ala	Gly	Cys	Thr	Tyr	Ile	Phe
				1445					1450					1455	
Glu	Lys	His	Ile	Gln	Ser	Thr	Arg	Glu	Ile	Phe	Leu	Phe	Glu	Arg	Tyr
				1460					1465					1470	
Gly	Gln	Val	Met	Glu	Tyr	Glu	Ser	Thr	Pro	Trp	Leu	Pro	Pro	Pro	Phe
				1475					1480					1485	
Thr	Ile	Ile	Tyr	His	Val	Ile	Trp	Leu	Phe	Lys	Leu	Ile	Lys	Ser	Ser
				1490					1495					1500	
Ser	Arg	Met	Phe	Glu	Arg	Lys	Asn	Leu	Phe	Asp	Gln	Ser	Leu	Lys	Leu
				1505					1510					1515	
Phe	Leu	Ser	Pro	Asp	Glu	Met	Glu	Lys	Val	His	Thr	Phe	Glu	Glu	Glu
				1525					1530					1535	
Ser	Val	Glu	Asp	Met	Lys	Arg	Glu	Thr	Glu	Lys	Lys	Asn	Leu	Ser	Ser
				1540					1545					1550	
Asn	Asp	Glu	Arg	Ile	His	Arg	Thr	Ala	Glu	Arg	Thr	Asp	Ala	Ile	Leu
				1555					1560					1565	
Asn	Arg	Val	Ser	His	Leu	Thr	Gln	Leu	Glu	Phe	Thr	Leu	Lys	Glu	Glu
				1570					1575					1580	
Ile	Arg	Glu	Leu	Glu	His	Lys	Met	Lys	Asn	Met	Asp	Ser	Arg	His	Lys
				1585					1590					1595	
Glu	Gln	Met	Asn	Leu	Met	Leu	Asp	Met	Asn	Lys	Lys	Leu	Gly	Lys	Phe
				1605					1610					1615	
Ile	Ser	Gly	Lys	Tyr	Lys	Arg	Gly	Ser	Phe	Gly	Gly	Ser	Gly	Ser	Asp
				1620					1625					1630	
Gly	Gly	Gly	Gly	Ser	Ser	Asp	Asn	Ser	Lys	Leu	Glu	Pro	Asn	Asn	Ser
				1635					1640					1645	
Val	Pro	Met	Ile	Thr	Val	Asp	Gly	Pro	Ser	Pro	Ile	Gly	Ser	Arg	Arg
				1650					1655					1660	
Thr	Ser	Gly	Gln	Tyr	Leu	Lys	Arg	Asp	Ser	Leu	Gln	Ala	Lys	Lys	Lys
				1665					1670					1675	
Ile	Thr	Glu	Asn	Arg	Arg	Ser	Ser	Leu	Glu	Gln	Pro	Lys	Ile	Pro	Ser
				1685					1690					1695	
Ile	Gln	Phe	Asn	Leu	Met	Glu	Asp	Gln	Asp	Glu	Ser	Ala	Ala	Glu	Ser
				1700					1705					1710	
Ala	Thr	Glu	Glu	Val	Ser	Ile	Ser	Ile	Pro	Val	Pro	Gln	Met	Arg	Val
				1715					1720					1725	
Arg	Gln	Val	Thr	Glu	Ser	Asp	Lys	Ser	Asp	Leu	Ser	Glu	Asp	Asp	Leu
				1730					1735					1740	
Ile	Thr	Arg	Glu	Asp	Ala	Pro	Pro	Thr	Ser	Ile	Asn	Leu	Pro	Arg	Gly
				1745					1750					1755	
Pro	Arg	Arg	His	Ala	Leu	Tyr	Ser	Thr	Ile	Ala	Asp	Ala	Ile	Glu	Thr

-22-

1765 1770 1775
 Glu Asp Asp Phe Tyr Ala Asp Ser Pro Val Pro Met Pro Met Thr Pro
 1780 1785 1790
 Val Gln Pro Ala Asp Gly Ser Phe Phe Gly Glu Asn Asp Ser Arg Tyr
 1795 1800 1805
 Gln Arg Asp Asp Ser Asp Tyr Glu
 1810 1815

<210> 14
 <211> 1387
 <212> PRT
 <213> C. Elegans

<400> 14
 Met Arg Lys Ser Arg Arg Val Arg Lys Leu Val Arg His Ala Ser Leu
 1 5 10 15
 Ile Glu Asn Ile Arg His Arg Thr Ser Ser Phe Leu Arg Leu Leu Asn
 20 25 30
 Ala Pro Arg Asn Ser Met Cys Asn Ala Asn Thr Val His Ser Ile Ser
 35 40 45
 Ser Phe Arg Ser Asp His Leu Ser Arg Lys Ser Thr His Lys Phe Leu
 50 55 60
 Asp Asn Pro Asn Leu Phe Ala Ile Glu Leu Thr Glu Lys Leu Ser Pro
 65 70 75 80
 Pro Trp Ile Glu Asn Thr Phe Glu Lys Arg Glu Cys Ile Arg Phe Ala
 85 90 95
 Ala Leu Pro Lys Asp Pro Glu Arg Cys Gly Cys Gly Arg Pro Leu Ser
 100 105 110
 Ala His Thr Pro Ala Ser Thr Phe Phe Ser Thr Leu Pro Val His Leu
 115 120 125
 Leu Glu Lys Glu Gln Gln Thr Trp Thr Ile Ala Asn Asn Thr Gln Thr
 130 135 140
 Ser Thr Thr Asp Ala Phe Gly Thr Ile Val Phe Gln Gly Gly Ala His
 145 150 155 160
 Ala His Lys Ala Gln Tyr Val Arg Leu Ser Tyr Asp Ser Glu Pro Leu
 165 170 175
 Asp Val Met Tyr Leu Met Glu Lys Val Trp Gly Leu Glu Ala Pro Arg
 180 185 190
 Leu Val Ile Thr Val His Gly Gly Met Ser Asn Phe Glu Leu Glu Glu
 195 200 205
 Arg Leu Gly Arg Leu Phe Arg Lys Gly Met Leu Lys Ala Ala Gln Thr
 210 215 220
 Thr Gly Ala Trp Ile Ile Thr Ser Gly Leu Asp Ser Gly Val Val Arg
 225 230 235 240
 His Val Ala Lys Ala Leu Asp Glu Ala Gly Ile Ser Ala Arg Met Arg
 245 250 255
 Ser Gln Ile Val Thr Ile Gly Ile Ala Pro Trp Gly Val Ile Lys Arg
 260 265 270
 Lys Glu Arg Leu Ile Arg Gln Asn Glu His Val Tyr Tyr Asp Val His
 275 280 285
 Ser Leu Ser Val Asn Ala Asn Val Gly Ile Leu Asn Asp Arg His Ser
 290 295 300
 Tyr Phe Leu Leu Ala Asp Asn Gly Thr Val Gly Arg Phe Gly Ala Asp
 305 310 315 320
 Leu His Leu Arg Gln Asn Leu Glu Asn His Ile Ala Thr Phe Gly Cys
 325 330 335
 Asn Gly Arg Lys Val Pro Val Val Cys Thr Leu Leu Glu Gly Gly Ile
 340 345 350
 Ser Ser Ile Asn Ala Ile His Asp Tyr Val Thr Met Lys Pro Asp Ile
 355 360 365

-23-

Pro	Ala	Ile	Val	Cys	Asp	Gly	Ser	Gly	Arg	Ala	Ala	Asp	Ile	Ile	Ser
370						375				380					
Phe	Ala	Ala	Arg	Tyr	Ile	Asn	Ser	Asp	Gly	Thr	Phe	Ala	Ala	Glu	Val
385					390					395					400
Gly	Glu	Lys	Leu	Arg	Asn	Leu	Ile	Lys	Met	Val	Phe	Pro	Glu	Thr	Asp
				405					410					415	
Gln	Glu	Glu	Met	Phe	Arg	Lys	Ile	Thr	Glu	Cys	Val	Ile	Arg	Asp	Asp
			420					425					430		
Leu	Leu	Arg	Ile	Phe	Arg	Tyr	Gly	Gln	Glu	Glu	Glu	Glu	Asp	Val	Asp
		435					440					445			
Phe	Val	Ile	Leu	Ser	Thr	Val	Leu	Gln	Lys	Gln	Asn	Leu	Pro	Pro	Asp
450						455				460					
Glu	Gln	Leu	Ala	Leu	Thr	Leu	Ser	Trp	Asn	Arg	Val	Asp	Leu	Ala	Lys
465					470					475					480
Ser	Cys	Leu	Phe	Ser	Asn	Gly	Arg	Lys	Trp	Ser	Ser	Asp	Val	Leu	Glu
				485					490					495	
Lys	Ala	Met	Asn	Asp	Ala	Leu	Tyr	Trp	Asp	Arg	Val	Asp	Phe	Val	Glu
			500					505					510		
Cys	Leu	Leu	Glu	Asn	Gly	Val	Ser	Met	Lys	Asn	Phe	Leu	Ser	Ile	Asn
		515					520					525			
Arg	Leu	Glu	Asn	Leu	Tyr	Asn	Met	Asp	Asp	Ile	Asn	Ser	Ala	His	Ser
530						535					540				
Val	Arg	Asn	Trp	Met	Glu	Asn	Phe	Asp	Ser	Met	Asp	Pro	His	Thr	Tyr
545					550					555					560
Leu	Thr	Ile	Pro	Met	Ile	Gly	Gln	Val	Val	Glu	Lys	Leu	Met	Gly	Asn
				565					570					575	
Ala	Phe	Gln	Leu	Tyr	Tyr	Thr	Ser	Arg	Ser	Phe	Lys	Gly	Lys	Tyr	Asp
			580					585					590		
Arg	Tyr	Lys	Arg	Ile	Asn	Gln	Ser	Ser	Tyr	Phe	His	Arg	Lys	Arg	Lys
		595					600					605			
Ile	Val	Gln	Lys	Glu	Leu	Phe	Lys	Lys	Lys	Ser	Asp	Asp	Gln	Ile	Asn
610						615					620				
Asp	Asn	Glu	Glu	Glu	Asp	Phe	Ser	Phe	Ala	Tyr	Pro	Phe	Asn	Asp	Leu
625					630					635					640
Leu	Ile	Trp	Ala	Val	Leu	Thr	Ser	Arg	His	Gly	Met	Ala	Glu	Cys	Met
				645					650					655	
Trp	Val	His	Gly	Glu	Asp	Ala	Met	Ala	Lys	Cys	Leu	Leu	Ala	Ile	Arg
				660				665					670		
Leu	Tyr	Lys	Ala	Thr	Ala	Lys	Ile	Ala	Glu	Asp	Glu	Tyr	Leu	Asp	Val
		675					680					685			
Glu	Glu	Ala	Lys	Arg	Leu	Phe	Asp	Asn	Ala	Val	Lys	Cys	Arg	Glu	Asp
690					695					700					
Ala	Ile	Glu	Leu	Leu	Asp	Gln	Cys	Tyr	Arg	Ala	Asp	His	Asp	Arg	Thr
705					710					715					720
Leu	Arg	Leu	Leu	Arg	Met	Glu	Leu	Pro	His	Trp	Gly	Asn	Asn	Asn	Cys
				725					730					735	
Leu	Ser	Leu	Ala	Val	Leu	Ala	Asn	Thr	Lys	Thr	Phe	Leu	Ala	His	Pro
			740					745					750		
Cys	Cys	Gln	Ile	Leu	Leu	Ala	Glu	Leu	Trp	His	Gly	Ser	Leu	Lys	Val
		755					760					765			
Arg	Ser	Gly	Ser	Asn	Val	Arg	Val	Leu	Thr	Ala	Leu	Ile	Cys	Pro	Pro
770						775					780				
Ala	Ile	Leu	Phe	Met	Ala	Tyr	Lys	Pro	Lys	His	Ser	Lys	Thr	Ala	Arg
785					790					795					800
Leu	Leu	Ser	Glu	Glu	Thr	Pro	Glu	Gln	Leu	Pro	Tyr	Pro	Arg	Glu	Ser
				805					810					815	
Ile	Thr	Ser	Thr	Thr	Ser	Asn	Arg	Tyr	Arg	Tyr	Ser	Lys	Gly	Pro	Glu
			820					825					830		
Glu	Gln	Lys	Glu	Thr	Leu	Leu	Glu	Lys	Gly	Ser	Tyr	Thr	Lys	Lys	Val
		835					840					845			

-24-

Thr	Ile	Ile	Ser	Ser	Arg	Lys	Asn	Ser	Gly	Val	Ala	Ser	Val	Tyr	Gly
	850					855					860				
Ser	Ala	Ser	Ser	Met	Met	Phe	Lys	Arg	Glu	Pro	Gln	Leu	Asn	Lys	Phe
865					870					875					880
Glu	Arg	Phe	Arg	Ala	Phe	Tyr	Ser	Ser	Pro	Ile	Thr	Lys	Phe	Trp	Ser
				885					890						895
Trp	Cys	Ile	Ala	Phe	Leu	Ile	Phe	Leu	Thr	Thr	Gln	Thr	Cys	Ile	Leu
			900					905						910	
Leu	Leu	Glu	Thr	Ser	Leu	Lys	Pro	Ser	Lys	Tyr	Glu	Trp	Ile	Thr	Phe
		915					920					925			
Ile	Tyr	Thr	Val	Thr	Leu	Ser	Val	Glu	His	Ile	Arg	Lys	Leu	Met	Thr
930						935					940				
Ser	Glu	Gly	Ser	Arg	Ile	Asn	Glu	Lys	Val	Lys	Val	Phe	Tyr	Ala	Lys
945					950					955					960
Trp	Tyr	Asn	Ile	Trp	Thr	Ser	Ala	Ala	Leu	Leu	Phe	Phe	Leu	Val	Gly
				965					970						975
Tyr	Gly	Phe	Arg	Leu	Val	Pro	Met	Tyr	Arg	His	Ser	Trp	Gly	Arg	Val
			980					985					990		
Leu	Leu	Ser	Phe	Ser	Asn	Val	Leu	Phe	Tyr	Met	Lys	Ile	Phe	Glu	Tyr
		995					1000					1005			
Leu	Ser	Val	His	Pro	Leu	Leu	Gly	Pro	Tyr	Ile	Gln	Met	Ala	Ala	Lys
		1010					1015				1020				
Met	Val	Trp	Ser	Met	Cys	Tyr	Ile	Cys	Val	Leu	Leu	Leu	Val	Pro	Leu
1025					1030					1035					104
Met	Ala	Phe	Gly	Val	Asn	Arg	Gln	Ala	Leu	Thr	Glu	Pro	Asn	Val	Lys
				1045					1050						1055
Asp	Trp	His	Trp	Leu	Leu	Val	Arg	Asn	Ile	Phe	Tyr	Lys	Pro	Tyr	Phe
			1060					1065					1070		
Met	Leu	Tyr	Gly	Glu	Val	Tyr	Ala	Gly	Glu	Ile	Asp	Thr	Cys	Gly	Asp
		1075					1080					1085			
Glu	Gly	Ile	Arg	Cys	Phe	Pro	Gly	Tyr	Phe	Ile	Pro	Pro	Leu	Leu	Met
	1090					1095					1100				
Val	Ile	Phe	Leu	Leu	Val	Ala	Asn	Ile	Leu	Leu	Leu	Asn	Leu	Leu	Ile
1105					1110					1115					112
Ala	Ile	Phe	Asn	Asn	Ile	Tyr	Asn	Asp	Ser	Ile	Glu	Lys	Ser	Lys	Glu
			1125						1130						1135
Ile	Trp	Leu	Phe	Gln	Arg	Tyr	Gln	Gln	Leu	Met	Glu	Tyr	His	Asp	Ser
			1140					1145						1150	
Pro	Phe	Leu	Pro	Pro	Pro	Phe	Ser	Ile	Phe	Ala	His	Val	Tyr	His	Phe
		1155					1160					1165			
Ile	Asp	Tyr	Leu	Tyr	Asn	Leu	Arg	Arg	Pro	Asp	Thr	Lys	Arg	Phe	Arg
	1170				1175						1180				
Ser	Glu	His	Ser	Ile	Lys	Leu	Ser	Val	Thr	Glu	Asp	Glu	Met	Lys	Arg
1185					1190					1195					120
Ile	Gln	Asp	Phe	Glu	Glu	Asp	Cys	Ile	Asp	Thr	Leu	Thr	Arg	Ile	Arg
			1205						1210						1215
Lys	Leu	Lys	Leu	Asn	Thr	Lys	Glu	Pro	Leu	Ser	Val	Thr	Asp	Leu	Thr
			1220					1225					1230		
Glu	Leu	Thr	Cys	Gln	Arg	Val	His	Asp	Leu	Met	Gln	Glu	Asn	Phe	Leu
		1235					1240					1245			
Leu	Lys	Ser	Arg	Val	Tyr	Asp	Ile	Glu	Thr	Lys	Ile	Asp	His	Ile	Ser
	1250					1255					1260				
Asn	Ser	Ser	Asp	Glu	Val	Val	Gln	Ile	Leu	Lys	Asn	Lys	Lys	Leu	Ser
1265					1270					1275					128
Gln	Asn	Phe	Ala	Ala	Ser	Ser	Leu	Ser	Leu	Pro	Asp	Thr	Ser	Ile	Glu
			1285						1290						1295
Val	Pro	Lys	Ile	Thr	Lys	Thr	Leu	Ile	Asp	Cys	His	Leu	Ser	Pro	Val
			1300					1305					1310		
Ser	Ile	Glu	Asp	Arg	Leu	Ala	Thr	Arg	Ser	Pro	Leu	Leu	Ala	Asn	Leu
		1315					1320						1325		

-25-

Gln Arg Asp His Thr Leu Arg Lys Leu Pro Thr Trp Glu Thr Ser Thr
 1330 1335 1340
 Ala Ser Thr Ser Ser Phe Glu Phe Val Phe Tyr Phe Thr Arg His Glu
 1345 1350 1355
 Gly Asn Glu Asn Lys Tyr Glu Phe Lys Lys Leu Glu Lys Gly Gly Phe
 1365 1370 1375
 Trp Arg Asn Asn Tyr Val Ile Ser Trp Arg Leu
 1380 1385

<210> 15
 <211> 1868
 <212> PRT
 <213> C. Elegans

<400> 15
 Met Asn Leu Cys Tyr Arg Arg His Arg Tyr Ala Ser Ser Pro Glu Val
 1 5 10 15
 Trp Cys Thr Met Glu Ser Asp Glu Leu Gly Val Thr Arg Tyr Leu Gln
 20 25 30
 Ser Lys Gly Gly Asp Gln Val Pro Pro Thr Ser Thr Thr Thr Gly Gly
 35 40 45
 Ala Gly Gly Asp Gly Asn Ala Val Pro Thr Thr Ser Gln Ala Gln Ala
 50 55 60
 Gln Thr Phe Asn Ser Gly Arg Gln Thr Thr Gly Met Ser Ser Gly Asp
 65 70 75 80
 Arg Leu Asn Glu Asp Val Ser Ala Thr Ala Asn Ser Ala Gln Leu Val
 85 90 95
 Leu Pro Thr Pro Leu Phe Asn Gln Met Arg Phe Thr Glu Ser Asn Met
 100 105 110
 Ser Leu Asn Arg His Asn Trp Val Arg Glu Thr Phe Thr Arg Arg Glu
 115 120 125
 Cys Ser Arg Phe Ile Ala Ser Ser Arg Asp Leu His Lys Cys Gly Cys
 130 135 140
 Gly Arg Thr Arg Asp Ala His Arg Asn Ile Pro Glu Leu Thr Ser Glu
 145 150 155 160
 Phe Leu Arg Gln Lys Arg Ser Val Ala Ala Leu Glu Gln Gln Arg Ser
 165 170 175
 Ile Ser Asn Val Asn Asp Asp Ile Asn Thr Gln Asn Met Tyr Thr Lys
 180 185 190
 Arg Gly Ala Asn Glu Lys Trp Ser Leu Arg Lys His Thr Val Ser Leu
 195 200 205
 Ala Thr Asn Ala Phe Gly Gln Val Glu Phe Gln Gly Gly Pro His Pro
 210 215 220
 Tyr Lys Ala Gln Tyr Val Arg Val Asn Phe Asp Thr Glu Pro Ala Tyr
 225 230 235 240
 Ile Met Ser Leu Phe Glu His Val Trp Gln Ile Ser Pro Pro Arg Leu
 245 250 255
 Ile Ile Thr Val His Gly Gly Thr Ser Asn Phe Asp Leu Gln Pro Lys
 260 265 270
 Leu Ala Arg Val Phe Arg Lys Gly Leu Leu Lys Ala Ala Ser Thr Thr
 275 280 285
 Gly Ala Trp Ile Ile Thr Ser Gly Cys Asp Thr Gly Val Val Lys His
 290 295 300
 Val Ala Ala Ala Leu Glu Gly Ala Gln Ser Ala Gln Arg Asn Lys Ile
 305 310 315 320
 Val Cys Ile Gly Ile Ala Pro Trp Gly Leu Lys Lys Arg Glu Asp
 325 330 335
 Phe Ile Gly Gln Asp Lys Thr Val Pro Tyr Tyr Pro Ser Ser Ser Lys
 340 345 350
 Gly Arg Phe Thr Gly Leu Asn Asn Arg His Ser Tyr Phe Leu Leu Val

355					360					365					
Asp	Asn	Gly	Thr	Val	Gly	Arg	Tyr	Gly	Ala	Glu	Val	Ile	Leu	Arg	Lys
370					375					380					
Arg	Leu	Glu	Met	Tyr	Ile	Ser	Gln	Lys	Gln	Lys	Ile	Phe	Gly	Gly	Thr
385					390					395					400
Arg	Ser	Val	Pro	Val	Val	Cys	Val	Val	Leu	Glu	Gly	Gly	Ser	Cys	Thr
				405					410					415	
Ile	Arg	Ser	Val	Leu	Asp	Tyr	Val	Thr	Asn	Val	Pro	Arg	Val	Pro	Val
			420					425					430		
Val	Val	Cys	Asp	Gly	Ser	Gly	Arg	Ala	Ala	Asp	Leu	Leu	Ala	Phe	Ala
		435					440					445			
His	Gln	Asn	Val	Thr	Glu	Asp	Gly	Leu	Leu	Pro	Asp	Asp	Ile	Arg	Arg
	450					455					460				
Gln	Val	Leu	Leu	Leu	Val	Glu	Thr	Thr	Phe	Gly	Cys	Ser	Glu	Ala	Ala
465					470					475					480
Ala	His	Arg	Leu	Leu	His	Glu	Leu	Thr	Val	Cys	Ala	Gln	His	Lys	Asn
				485					490					495	
Leu	Leu	Thr	Ile	Phe	Arg	Leu	Gly	Glu	Gln	Gly	Glu	His	Asp	Val	Asp
			500					505					510		
His	Ala	Ile	Leu	Thr	Ala	Leu	Leu	Lys	Gly	Gln	Asn	Leu	Ser	Ala	Ala
		515					520					525			
Asp	Gln	Leu	Ala	Leu	Ala	Leu	Ala	Trp	Asn	Arg	Val	Asp	Ile	Ala	Arg
				535							540				
Ser	Asp	Val	Phe	Ala	Met	Gly	His	Glu	Trp	Pro	Gln	Ala	Ala	Leu	His
545					550					555					560
Asn	Ala	Met	Met	Glu	Ala	Leu	Ile	His	Asp	Arg	Val	Asp	Phe	Val	Arg
				565					570					575	
Leu	Leu	Leu	Glu	Gln	Gly	Ile	Asn	Met	Gln	Lys	Phe	Leu	Thr	Ile	Ser
			580				585						590		
Arg	Leu	Asp	Glu	Leu	Tyr	Asn	Thr	Asp	Lys	Gly	Pro	Pro	Asn	Thr	Leu
		595					600					605			
Phe	Tyr	Ile	Val	Arg	Asp	Val	Val	Arg	Val	Arg	Gln	Gly	Tyr	Arg	Phe
610					615					620					
Lys	Leu	Pro	Asp	Ile	Gly	Leu	Val	Ile	Glu	Lys	Leu	Met	Gly	Asn	Ser
625					630					635					640
Tyr	Gln	Cys	Ser	Tyr	Thr	Thr	Ser	Glu	Phe	Arg	Asp	Lys	Tyr	Lys	Gln
				645					650					655	
Arg	Met	Lys	Arg	Val	Lys	His	Ala	Gln	Lys	Lys	Ala	Met	Gly	Val	Phe
			660					665					670		
Ser	Ser	Arg	Pro	Ser	Arg	Thr	Gly	Ser	Gly	Ile	Ala	Ser	Arg	Gln	Ser
		675					680					685			
Thr	Glu	Gly	Met	Gly	Gly	Val	Gly	Gly	Gly	Ser	Ser	Val	Ala	Gly	Val
				695								700			
Phe	Gly	Asn	Ser	Phe	Gly	Asn	Gln	Asp	Pro	Pro	Leu	Asp	Pro	His	Val
705				710						715					720
Asn	Arg	Ser	Ala	Leu	Ser	Gly	Ser	Arg	Ala	Leu	Ser	Asn	His	Ile	Leu
				725					730					735	
Trp	Arg	Ser	Ala	Phe	Arg	Gly	Asn	Phe	Pro	Ala	Asn	Pro	Met	Arg	Pro
			740				745					750			
Pro	Asn	Leu	Gly	Asp	Ser	Arg	Asp	Cys	Gly	Ser	Glu	Phe	Asp	Glu	Glu
		755					760					765			
Leu	Ser	Leu	Thr	Ser	Ala	Ser	Asp	Gly	Ser	Gln	Thr	Glu	Pro	Asp	Phe
				775							780				
Arg	Tyr	Pro	Tyr	Ser	Glu	Leu	Met	Ile	Trp	Ala	Val	Leu	Thr	Lys	Arg
785				790						795					800
Gln	Asp	Met	Ala	Met	Cys	Met	Trp	Gln	His	Gly	Glu	Glu	Ala	Met	Ala
				805					810					815	
Lys	Ala	Leu	Val	Ala	Cys	Arg	Leu	Tyr	Lys	Ser	Leu	Ala	Thr	Glu	Ala
			820					825					830		
Ala	Glu	Asp	Tyr	Leu	Glu	Val	Glu	Ile	Cys	Glu	Glu	Leu	Lys	Lys	Tyr

-27-

Ala	Glu	Glu	Phe	Arg	Ile	Leu	Ser	Leu	Glu	Leu	Leu	Asp	His	Cys	Tyr
850						855					860				
His	Val	Asp	Asp	Ala	Gln	Thr	Leu	Gln	Leu	Leu	Thr	Tyr	Glu	Leu	Ser
865					870					875					880
Asn	Trp	Ser	Asn	Glu	Thr	Cys	Leu	Ala	Leu	Ala	Val	Ile	Val	Asn	Asn
			885						890					895	
Lys	His	Phe	Leu	Ala	His	Pro	Cys	Cys	Gln	Ile	Leu	Leu	Ala	Asp	Leu
		900					905					910			
Trp	His	Gly	Gly	Leu	Arg	Met	Arg	Thr	His	Ser	Asn	Ile	Lys	Val	Val
	915					920						925			
Leu	Gly	Leu	Ile	Cys	Pro	Pro	Phe	Ile	Gln	Met	Leu	Glu	Phe	Lys	Thr
930					935						940				
Arg	Glu	Glu	Leu	Leu	Asn	Gln	Pro	Gln	Thr	Ala	Ala	Glu	His	Gln	Asn
945					950					955					960
Asp	Met	Asn	Tyr	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser
			965						970						975
Ser	Ser	Ser	Ser	Ser	Asp	Ser	Ser	Ser	Ser	Phe	Glu	Asp	Asp	Asp	Glu
		980					985						990		
Asn	Asn	Ala	His	Asn	His	Asp	Gln	Lys	Arg	Thr	Arg	Lys	Thr	Ser	Gln
	995						1000					1005			
Gly	Ser	Ala	Gln	Ser	Leu	Asn	Ile	Thr	Ser	Leu	Phe	His	Ser	Arg	Arg
1010					1015						1020				
Arg	Lys	Ala	Lys	Lys	Asn	Glu	Lys	Cys	Asp	Arg	Glu	Thr	Asp	Ala	Ser
1025					1030					1035					104
Ala	Cys	Glu	Ala	Gly	Asn	Arg	Gln	Ile	Gln	Asn	Gly	Gly	Leu	Thr	Ala
			1045						1050					1055	
Glu	Tyr	Gly	Thr	Phe	Gly	Glu	Ser	Asn	Gly	Val	Ser	Pro	Pro	Pro	Pro
	1060						1065					1070			
Tyr	Met	Arg	Ala	Asn	Ser	Arg	Ser	Arg	Tyr	Asn	Asn	Arg	Ser	Asp	Met
1075						1080						1085			
Ser	Lys	Thr	Ser	Ser	Val	Ile	Phe	Gly	Ser	Asp	Pro	Asn	Leu	Ser	Lys
1090					1095						1100				
Leu	Gln	Lys	Ser	Asn	Ile	Thr	Ser	Thr	Asp	Arg	Pro	Asn	Pro	Met	Glu
1105				1110						1115					112
Gln	Phe	Gln	Gly	Thr	Arg	Lys	Ile	Lys	Met	Arg	Arg	Arg	Phe	Tyr	Glu
			1125						1130					1135	
Phe	Tyr	Ser	Ala	Pro	Ile	Ser	Thr	Phe	Trp	Ser	Trp	Thr	Ile	Ser	Phe
	1140						1145					1150			
Ile	Leu	Phe	Ile	Thr	Phe	Phe	Thr	Tyr	Thr	Leu	Leu	Val	Lys	Thr	Pro
	1155						1160					1165			
Pro	Arg	Pro	Thr	Val	Ile	Glu	Tyr	Ile	Leu	Ile	Ala	Tyr	Val	Ala	Ala
1170					1175						1180				
Phe	Gly	Leu	Glu	Gln	Val	Arg	Lys	Ile	Ile	Met	Ser	Asp	Ala	Lys	Pro
1185				1190						1195					120
Phe	Tyr	Glu	Lys	Ile	Arg	Thr	Tyr	Val	Cys	Ser	Phe	Trp	Asn	Cys	Val
			1205						1210					1215	
Thr	Ile	Leu	Ala	Ile	Phe	Tyr	Ile	Val	Gly	Phe	Phe	Met	Arg	Cys	
	1220						1225					1230			
Phe	Gly	Ser	Val	Ala	Tyr	Gly	Arg	Val	Ile	Leu	Ala	Cys	Asp	Ser	Val
	1235						1240					1245			
Leu	Trp	Thr	Met	Lys	Leu	Leu	Asp	Tyr	Met	Ser	Val	His	Pro	Lys	Leu
1250					1255						1260				
Gly	Pro	Tyr	Val	Thr	Met	Ala	Gly	Lys	Met	Ile	Gln	Asn	Met	Ser	Tyr
1265				1270						1275					128
Ile	Ile	Val	Met	Leu	Val	Val	Thr	Leu	Leu	Ser	Phe	Gly	Leu	Ala	Arg
			1285						1290					1295	
Gln	Ser	Ile	Thr	Tyr	Pro	Asp	Glu	Thr	Trp	His	Trp	Ile	Leu	Val	Arg
	1300						1305					1310			
Asn	Ile	Phe	Leu	Lys	Pro	Tyr	Phe	Met	Leu	Tyr	Gly	Glu	Val	Tyr	Ala

1315	1320	1325
Asp Glu Ile Asp Thr Cys Gly Asp Glu Ala Trp Asp Gln His Leu Glu		
1330	1335	1340
Asn Gly Gly Pro Val Ile Leu Gly Asn Gly Thr Thr Gly Leu Ser Cys		
1345	1350	1355
Val Pro Gly Tyr Trp Ile Pro Pro Leu Leu Met Thr Phe Phe Leu Leu		
1365	1370	1375
Ile Ala Asn Ile Leu Leu Met Ser Met Leu Ile Ala Ile Phe Asn His		
1380	1385	1390
Ile Phe Asp Ala Thr Asp Glu Met Ser Gln Gln Ile Trp Leu Phe Gln		
1395	1400	1405
Arg Tyr Lys Gln Val Met Glu Tyr Glu Ser Thr Pro Phe Leu Pro Pro		
1410	1415	1420
Pro Leu Thr Pro Leu Tyr His Gly Val Leu Ile Leu Gln Phe Val Arg		
1425	1430	1435
Thr Arg Leu Ser Cys Ser Lys Ser Gln Glu Arg Asn Pro Ile Leu Leu		
1445	1450	1455
Leu Lys Ile Ala Glu Leu Phe Leu Asp Asn Asp Gln Ile Glu Lys Leu		
1460	1465	1470
His Asp Phe Glu Glu Asp Cys Met Glu Asp Leu Ala Arg Gln Lys Leu		
1475	1480	1485
Asn Glu Lys Asn Thr Ser Asn Glu Gln Arg Ile Leu Arg Ala Asp Ile		
1490	1495	1500
Arg Thr Asp Gln Ile Leu Asn Arg Leu Ile Asp Leu Gln Ala Lys Glu		
1505	1510	1515
Ser Met Gly Arg Asp Val Ile Asn Asp Val Glu Ser Arg Leu Ala Ser		
1525	1530	1535
Val Glu Lys Ala Gln Asn Glu Ile Leu Glu Cys Val Arg Ala Leu Leu		
1540	1545	1550
Asn Gln Asn Asn Ala Pro Thr Ala Ile Gly Arg Cys Phe Ser Pro Ser		
1555	1560	1565
Pro Asp Pro Leu Val Glu Thr Ala Asn Gly Thr Pro Gly Pro Leu Leu		
1570	1575	1580
Leu Lys Leu Pro Gly Thr Asp Pro Ile Leu Glu Glu Lys Asp His Asp		
1585	1590	1595
Ser Gly Glu Asn Ser Asn Ser Leu Pro Pro Gly Arg Ile Arg Arg Asn		
1605	1610	1615
Arg Thr Ala Thr Ile Cys Gly Gly Tyr Val Ser Glu Glu Arg Asn Met		
1620	1625	1630
Met Leu Leu Ser Pro Lys Pro Ser Asp Val Ser Gly Ile Pro Gln Gln		
1635	1640	1645
Arg Leu Met Ser Val Thr Ser Met Asp Pro Leu Pro Leu Pro Leu Ala		
1650	1655	1660
Lys Leu Ser Thr Met Ser Ile Arg Arg Arg His Glu Glu Tyr Thr Ser		
1665	1670	1675
Ile Thr Asp Ser Ile Ala Ile Arg His Pro Glu Arg Arg Ile Arg Asn		
1685	1690	1695
Asn Arg Ser Asn Ser Ser Glu His Asp Glu Ser Ala Val Asp Ser Glu		
1700	1705	1710
Gly Gly Gly Asn Val Thr Ser Ser Pro Arg Lys Arg Ser Thr Arg Asp		
1715	1720	1725
Leu Arg Met Thr Pro Ser Ser Gln Val Glu Glu Ser Thr Ser Arg Asp		
1730	1735	1740
Gln Ile Phe Glu Ile Asp His Pro Glu His Glu Glu Asp Glu Ala Gln		
1745	1750	1755
Ala Asp Cys Glu Leu Thr Asp Val Ile Thr Glu Glu Glu Asp Glu Glu		
1765	1770	1775
Glu Asp Asp Glu Glu Asp Asp Ser His Glu Arg His His Ile His Pro		
1780	1785	1790
Arg Arg Lys Ser Ser Arg Gln Asn Arg Gln Pro Ser His Thr Leu Glu		

-29-

1795 1800 1805
 Thr Asp Leu Ser Glu Gly Glu Glu Val Asp Pro Leu Asp Val Leu Lys
 1810 1815 1820
 Met Lys Glu Leu Pro Ile Ile His Gln Ile Leu Asn Glu Glu Glu Gln
 1825 1830 1835 184
 Ala Gly Ala Pro His Ser Thr Pro Val Ile Ala Ser Pro Ser Ser Ser
 1845 1850 1855
 Arg Ala Asp Leu Thr Ser Gln Lys Cys Ser Asp Val
 1860 1865

<210> 16
 <211> 489
 <212> DNA
 <213> Mus Musculus

<400> 16
 ccctgaaaga ctgcacttct gctgctagcg ctggagctga gttagttttg agaaggtttc 60
 ccggggctgt ccttggttcgg tggcccgtgc caccgcctcc ggagacgctt tccgatagat 120
 ggctgcaggc cgcgagggtg gaggaggagc cgctgccctt ccggagtcctg ccccgtagagg 180
 agaatgtccc agaaatcctg gatagagagc actttgacca agagggagtg tgtatatatt 240
 ataccaagct ccaaagacc cccacagatgt cttccaggat gtcagatttg tcagcaactt 300
 gtcagatggt tctgtggtcg tttggtcaag caacatgcat gctttactgc aagtcttgcc 360
 atgaaatact cagatgtgaa attgggtgaa cactttaacc aggcaataga agaatggtct 420
 gtggaaaagc acacggagca gagcccaaca gatgcttatg gagtcatcaa ttttcaaggg 480
 ggttctcat 489

<210> 17
 <211> 102
 <212> PRT
 <213> Mus Musculus

<400> 17
 Met Ser Gln Lys Ser Trp Ile Glu Ser Thr Leu Thr Lys Arg Glu Cys
 1 5 10 15
 Val Tyr Ile Ile Pro Ser Ser Lys Asp Pro His Arg Cys Leu Pro Gly
 20 25 30
 Cys Gln Ile Cys Gln Gln Leu Val Arg Cys Phe Cys Gly Arg Leu Val
 35 40 45
 Lys Gln His Ala Cys Phe Thr Ala Ser Leu Ala Met Lys Tyr Ser Asp
 50 55 60
 Val Lys Leu Gly Glu His Phe Asn Gln Ala Ile Glu Glu Trp Ser Val
 65 70 75 80
 Glu Lys His Thr Glu Gln Ser Pro Thr Asp Ala Tyr Gly Val Ile Asn
 85 90 95
 Phe Gln Gly Gly Ser His
 100

<210> 18
 <211> 410
 <212> DNA
 <213> Homo Sapiens

<220>
 <221> unsure
 <222> (6)...(6)
 <221> unsure
 <222> (58)...(58)
 <221> unsure

-30-

<222> (89)...(89)

<221> unsure

<222> (406)...(406)

<400> 18

gccgcnggag	cctgagcgga	gggtgtgcgc	agcctcgcca	gcggggggccc	cgggctgngc	60
cattgcctca	ctgagccagc	gcctgcctnc	tacctcgccg	acagctggaa	ccagtgcgac	120
ctagtggctc	tcacctgctt	cctcctgggc	gtgggctgcc	ggctgacccc	gggtttgtac	180
cacctgggccc	gcactgtcct	ctgcatcgac	ttcatgggtt	tcacggtgcg	gctgcttcac	240
atcttcacgg	tcaacaaaca	gctggggccc	aagatcgta	tcgtgagcaa	gatgatgaag	300
gacgtgttct	tcttcctctt	cttcctcggc	gtgtggctgg	tagctatggg	ttgggccacg	360
gaggggttcc	tgaggccacg	ggacagtgc	ttcccaagta	tcctgncgcc		410

<210> 19

<211> 131

<212> PRT

<213> Homo Sapiens

<220>

<221> UNSURE

<222> (15)...(15)

<223> UNKNOWN

<221> UNSURE

<222> (25)...(25)

<223> UNKNOWN

<221> UNSURE

<222> (131)...(131)

<223> UNKNOWN

<400> 19

Ala	Glu	Gly	Val	Arg	Ser	Leu	Ala	Ser	Gly	Gly	Pro	Gly	Leu	Xaa	His
1				5					10					15	
Cys	Leu	Thr	Glu	Pro	Ala	Pro	Ala	Xaa	Tyr	Leu	Ala	Asp	Ser	Trp	Asn
			20					25					30		
Gln	Cys	Asp	Leu	Val	Ala	Leu	Thr	Cys	Phe	Leu	Leu	Gly	Val	Gly	Cys
			35				40					45			
Arg	Leu	Thr	Pro	Gly	Leu	Tyr	His	Leu	Gly	Arg	Thr	Val	Leu	Cys	Ile
			50				55				60				
Asp	Phe	Met	Val	Phe	Thr	Val	Arg	Leu	Leu	His	Ile	Phe	Thr	Val	Asn
			65			70				75				80	
Lys	Gln	Leu	Gly	Pro	Lys	Ile	Val	Ile	Val	Ser	Lys	Met	Met	Lys	Asp
			85						90					95	
Val	Phe	Phe	Phe	Leu	Phe	Phe	Leu	Gly	Val	Trp	Leu	Val	Ala	Met	Gly
			100					105					110		
Trp	Ala	Thr	Glu	Gly	Phe	Leu	Arg	Pro	Arg	Asp	Ser	Asp	Phe	Pro	Ser
			115				120					125			
Ile	Leu	Xaa													
			130												

<210> 20

<211> 389

<212> DNA

<213> Homo Sapiens

<400> 20

caaatTTTTT	gttagtacac	catctcatcc	aaattgcaaa	agtcacatgg	aaactggaac	60
caaagatcaa	gaaactgttt	gctctaaagc	tacagaagga	gataatacag	aatttgagac	120

-31-

attttagtagga	cacagagata	gcatggattt	acagagggtt	aaagaaacat	caaacaagat	180
aaaaatacta	tccaataaca	atacttctga	aaacactttg	aaacgagtga	gttctcttgc	240
tggttttact	gactgtcaca	gaacttccat	tcctgttcat	tcaaaacgag	aaaagatcag	300
tagaaggcca	tctaccgaag	acactcatga	agtagattcc	aaagcagctt	taataccggt	360
ttgtagattt	caactaaaca	gatatatat				389

<210> 21
 <211> 415
 <212> DNA
 <213> Homo Sapiens

<400> 21						
atttctagtt	tttcaaattt	gccagtcctt	ttgaatagta	tctccttctt	ttctcatggt	60
ttatatTTaa	aactttttta	tgtccatcat	cacttttaac	atacttattt	tgtcatctat	120
aaccaataat	tccactatct	tatcagaaat	caaataccgt	ttatgtaagt	tgactcccat	180
gagttctaaa	ttgccattgt	gaggtcatct	tcggttaggc	tttaatttgt	tgcaaagtgt	240
tgcagctcag	ggtcaggaag	agtcctcca	gaaaggagga	tttgttactg	tgaatctctt	300
tgTTaaactaa	cctctttccc	cactgaaata	acttttttca	ataacatgat	tttaacaaca	360
taatctctct	atgccagaac	agatatatat	gaatgtaagt	caatatTTtc	ttgag	415

<210> 22
 <211> 405
 <212> DNA
 <213> Mus Musculus

<400> 22						
ttattatggc	ttatcatgaa	aaaccagtcc	tgccctctcc	tcttatcatc	ctcagccata	60
tagtttctact	gttttgctgt	gtatgcaaaa	gaagaaagaa	agataagact	tccgatgggc	120
caaaactttt	cttaacagaa	gaagatcaaa	agaaactcca	tgattttgaa	gagcagtgtg	180
ttgagatgta	ctttgatgag	aaagatgaca	aattcaattc	tgggagtga	gagagaatcc	240
gggtcacttt	tgaaagagtg	gagcagatga	gcattcagat	taaagaagtt	ggagatcggtg	300
tcaactacat	aaaaagatca	ttacagtctt	tagattctca	aattggtcat	ctgcaagatc	360
tctcagccct	aacagtagat	acattgaaaa	cacttacagc	ccaga		405

<210> 23
 <211> 5117
 <212> DNA
 <213> Homo Sapiens

<220>
 <221> unsure
 <222> (2382) ... (2382)
 <223> unknown

<221> unsure
 <222> (4664) ... (4664)
 <223> unknown

<221> unsure
 <222> (4682) ... (4682)
 <223> unknown

<221> unsure
 <222> (4702) ... (4702)
 <223> unknown

<221> unsure
 <222> (5038) ... (5039)
 <223> unknown

<221> unsure
 <222> (5056)... (5056)
 <223> unknown

<221> unsure
 <222> (5071)... (5072)

<400> 23

gatggcaaca	tggtgaagaa	tcaatggcta	aagcattagt	tgctgtgaag	atctatcggt	60
caatggcata	tgaagcaaa	cagagtgaac	tggtagatga	tacttcagaa	gaactaaaac	120
agtattccaa	tgattttggt	cagttggccg	ttgaattatt	agaacagtcc	ttcagacaag	180
atgaaaccaa	ggctatgaaa	ttgctcactt	atgaactgaa	gaactggagt	aattcaacct	240
gccttaagtt	agcagtttct	tcaagactta	gaccttttgt	agctcacacc	tgtaacacaa	300
tggtgttatc	tgatatgtgg	atgggaaggc	tgaaatagag	gaaaaattcc	tggtacaagg	360
tcatactaag	catttttagtt	ccacctgcc	tattgctgtt	agagtataaa	actaaggctg	420
aaatgtccca	tatcccacaa	tctcaagatg	ctcatcagat	gacaatggat	gacagcgaaa	480
acaactttca	gaacataaca	gaagagatcc	ccatggaagt	gtttaaagaa	gtacggattt	540
tggaatagta	tgaaggaaa	aatgagatgg	agatacaaat	gaaatcaaaa	aagcttccaa	600
ttacgcgaaa	gttttatgcc	ttttatcatg	caccaattgt	aaaattctgg	tttaacacgt	660
tggtcatatt	aggatttctg	atgctttata	catttgtggt	tcttgtacaa	atggaacagt	720
taccttcagt	tcaagaatgg	attgttattg	cttatatttt	tacttatgcc	attgagaaag	780
tccgtgagat	ctttatgtct	gaagctggga	aagtaaacca	gaagattaaa	gtatggttta	840
gtgattactt	caacatcagt	gatacaattg	ccataatttc	tttcttcatt	ggatttggac	900
taagatttgg	agcaaaatgg	aactttgcaa	atgcatatga	taatcatggt	tttgtggctg	960
gaagattaat	ttactgtctt	aacataatat	tttggtatgt	gcgtttgcta	gattttctag	1020
ctgtaaatca	acaggcagga	ccttatgtaa	tgatgattgg	aaaaatggtg	gccaatatgt	1080
tctacattgt	agtgattatg	gctcttgat	tacttagttt	tggtgttccc	agaaaggcaa	1140
tactttatcc	tcatgaagca	ccatcttgga	ctcttgctaa	agatatagtt	tttcacccat	1200
actggatgat	ttttggtgaa	gtttatgcat	acgaaattga	tgtgtgtgca	aatgattctg	1260
ttatccctca	aatctgtggt	cctgggacgt	ggttgactcc	atttcttcaa	gcagtctacc	1320
tctttgtaca	gtatatcatt	atggttaatt	ttcttattgc	atttttcaac	aatgtgtatt	1380
tacaagtga	ggcaatttcc	aatattgtat	ggaagtacca	gcgttatcat	tttattatgg	1440
cttatcatga	gaaaccagtt	ctgcctcctc	cacttatcat	tcttagccat	atagtttctc	1500
tgttttgctg	catatgtaag	agaagaaa	aagataagac	ttccgatgga	ccaaaacttt	1560
tcttaacaga	agaagatcaa	aagaaaactc	atgattttga	agagcagtg	gttgaaatgt	1620
atttcaatga	aaaagatgac	aaatttctt	ctgggagtga	agagagaatt	cgtgtcactt	1680
ttgaaagagt	ggaacagatg	tgcatcaga	ttaaagaagt	tgagatcgt	gtcaactaca	1740
taaaaagatc	attacaatca	ttagattctc	aaattggcca	tttgcaagat	ctttcagccc	1800
tgacggtaga	tacattaaaa	acactcactg	cccagaaagc	gtcggaagct	agcaaagttc	1860
ataatgaaat	cacacagaaa	ctgagcattt	ccaaacactt	ggctcaaaac	cttattgatg	1920
atggtcctgt	aagaccttct	gtatggaaaa	agcatggtgt	tgtaaataca	cttagctcct	1980
ctcttctcca	aggtgatctt	gaaagtata	atccttttca	ttgtaattat	ttaatgaaag	2040
atgacaaa	tccccagtg	aatatatatt	gtcaagactt	acctgcagta	ccccagagaa	2100
aagaatttaa	ttttccagag	gctggttctt	cttctgtgtc	cttattccca	agtgtgtttt	2160
cccctccaga	actgcgacag	agactacatg	gggtagaact	cttaaaaaata	tttaataaaa	2220
atcaaaaatt	aggcagttca	tctactagca	taccacatct	gtcatcccca	ccaaccaa	2280
tttttgtag	tacaccatct	cagccaagtt	gcaaaagcca	cttggaact	ggaaccaaag	2340
atcaagaaac	tgtttgcctc	aaagctacag	aaggagataa	tncagaattt	ggagcatttg	2400
taggacacag	agatagcatg	gatttacaga	ggtttaaga	aacatcaaac	aagataaaaa	2460
tactatccaa	taacaatact	tctgaaaaca	ctttgaaacg	agtgagttct	cttgctggat	2520
ttactgactg	tcacagaact	tccattcctg	ttcattcaaa	acaagcagaa	aaaatcagta	2580
gaaggccatc	taccgaagac	actcatgaag	tagattccaa	agcagcttta	ataccggatt	2640
ggttacaaga	tagaccatca	aacagagaaa	tgccatctga	agaaggaaca	ttaaatggtc	2700
tcacttctcc	atttaagcca	gctatggata	caaattacta	ttattcagct	gtggaaagaa	2760
ataacttgat	gaggttatca	cagagcattc	catttacacc	tgtgcctcca	agaggggagc	2820
ctgtcacagt	gtatcgtttg	gaagagagtt	cacccaacat	actaaataac	agcatgtcct	2880
cttggtcaca	actagcctc	tgtgccaaaa	tagagttttt	aagcaaaagag	gagatgggag	2940
gaggtttacg	aagagctgtc	aaagtacagt	gtacctggtc	agaacatgat	atcctcaaat	3000
cagggcatct	ttatattatc	aaatcttttc	ttccagaggt	ggttaataca	tggtcaagta	3060
tttataaaga	agatacagtt	ctgcatctct	gtctgagaga	aattcaaaaa	cagagagcag	3120

-33-

```

cacaaaagct tacgtttgcc tttaatcaaa tgaaacccaa atccatacca tattctccaa 3180
gggttccttga agttttcctg ctgtattgcc attcagcagg acagtgggtt gctgtggaag 3240
aatgtatgac tggagaatth agaaaataca acaataataa tggagatgag attattccaa 3300
ctaatactct ggaagagatc atgctagcct ttagccactg gacttacgaa tatacaagag 3360
gggagttact ggtacttgat ttgcaagggt ttggtgaaaa tttgactgac ccatctgtga 3420
taaaagcaga agaaaagaga tcctgtgata tgggttttgg cccagcaaat ctaggagaag 3480
atgcaattaa aaacttcaga gcaaaacatc actgtaattc ttgctgtaga aagcttaaac 3540
ttccagatct gaagaggaat gattatacgc ctgataaaat tatatttctt caggatgagc 3600
cttcagatth gaatcttcag cctggaaatt ccaccaaaga atcagaatca gctaattctg 3660
ttcgtctgat gttataatat taatattact gaatcattgg ttttgccctg acctcacaga 3720
aatgttactg tgtcacttht ccctcgggag gaaattgttt ggtaatatag aaaggtgtat 3780
gcaagttgaa tttgtgact ccagcacagt taaaaggcca atattctttt gacctgatta 3840
atcagtcaga aagtccttat aggatagagc tggcagctga gaaattthaa aggtaatga 3900
taattagtat tttaacttht ttaaaagggt cttgtatag cagaggatct catttgactt 3960
tgthttgatg aggtgatgc cctctcttat gtggtacaat accattaacc aaaggtaggt 4020
gtccatgcag attttattgg cagctgttht attgccattc aactagggaa atgaagaaat 4080
cacgcagcct tttggttaaa tggcagtcac aattttcctc agtgatttht gtgtgttcag 4140
tgatgatata actggttccc aactagatgc ttgttgacca cgggaaggga aatgacttgt 4200
tctaattcta ggttcacaga ggtatgagaa gcctgaactg aagaccattt tcaagaggga 4260
cggtatttat gaatcagggt taggctccat atttaaagat agagccagtt tttttthtaa 4320
atagaaccca aattgtgtaa aaatgttaat tgggttttht aaacattgtt ttatcaagtc 4380
actgttaagt agaagaaagc catggtaaac tgatacataa cctaaattat aaaagcagaa 4440
acctaactca ctctgcaagg gaagttacct tttgaggaaa gttaaagtac ttttttccct 4500
atctgtatct atagcaacaa cccagaactt acaaacttct ccaaagattt tattgattgt 4560
tatatcaaat cagaatgtaa acatgaactc ttgcataat ttaaaattgt gttggaacat 4620
ttgaacatga atgctgtttg ggtacttaag aaattrattc agtnngatta tcattatgtg 4680
anactggcag attgcagtgc anccttatgc caataaaatg taatttaaca gccccagata 4740
ttgttgaata ttcaacaata acaagaaaag cttttcatct aagttttatg ctttaatttt 4800
ttttcttttt ttttcttttt cttttgtttc cttggtacta attttaattt ttatttgtaa 4860
gggagcagta taaagcttat ttgtatttag tagtgtatct catagataca gacaaggcaa 4920
gagatgataa gctgtthtaa tagtgtthaa tattgattgg ggggtggggag aaagaaaaag 4980
tgtattactt aaagatacta tatacgttht gtatatcatt aaatctthaa aagaaatnna 5040
ataaatttat tgthtncaaa aaaaaaaccc nntaaaaaaa aaagggcggc ccctctagag 5100
gatccctoga ggggccc 5117

```

<210> 24
 <211> 1224
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> UNSURE
 <222> (794)... (794)
 <223> UNKNOWN

<400> 24
 Trp Gln His Gly Glu Ser Met Ala Lys Ala Leu Val Ala Cys Lys
 1 5 10 15
 Ile Tyr Arg Ser Met Ala Tyr Glu Ala Lys Gln Ser Asp Leu Val Asp
 20 25 30
 Asp Thr Ser Glu Glu Leu Lys Gln Tyr Ser Asn Asp Phe Gly Gln Leu
 35 40 45
 Ala Val Glu Leu Leu Glu Gln Ser Phe Arg Gln Asp Glu Thr Met Ala
 50 55 60
 Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ser Thr Cys
 65 70 75 80
 Leu Lys Leu Ala Val Ser Ser Arg Leu Arg Pro Phe Val Ala His Thr
 85 90 95

-34-

Cys	Thr	Gln	Met	Leu	Leu	Ser	Asp	Met	Trp	Met	Gly	Arg	Leu	Asn	Met
			100					105					110		
Arg	Lys	Asn	Ser	Trp	Tyr	Lys	Val	Ile	Leu	Ser	Ile	Leu	Val	Pro	Pro
		115					120					125			
Ala	Ile	Leu	Leu	Leu	Glu	Tyr	Lys	Thr	Lys	Ala	Glu	Met	Ser	His	Ile
	130					135					140				
Pro	Gln	Ser	Gln	Asp	Ala	His	Gln	Met	Thr	Met	Asp	Asp	Ser	Glu	Asn
145					150					155				160	
Asn	Phe	Gln	Asn	Ile	Thr	Glu	Glu	Ile	Pro	Met	Glu	Val	Phe	Lys	Glu
			165						170					175	
Val	Arg	Ile	Leu	Asp	Ser	Asn	Glu	Gly	Lys	Asn	Glu	Met	Glu	Ile	Gln
		180						185						190	
Met	Lys	Ser	Lys	Lys	Leu	Pro	Ile	Thr	Arg	Lys	Phe	Tyr	Ala	Phe	Tyr
		195					200						205		
His	Ala	Pro	Ile	Val	Lys	Phe	Trp	Phe	Asn	Thr	Leu	Ala	Tyr	Leu	Gly
	210					215					220				
Phe	Leu	Met	Leu	Tyr	Thr	Phe	Val	Val	Leu	Val	Gln	Met	Glu	Gln	Leu
225					230					235					240
Pro	Ser	Val	Gln	Glu	Trp	Ile	Val	Ile	Ala	Tyr	Ile	Phe	Thr	Tyr	Ala
				245					250					255	
Ile	Glu	Lys	Val	Arg	Glu	Ile	Phe	Met	Ser	Glu	Ala	Gly	Lys	Val	Asn
			260					265						270	
Gln	Lys	Ile	Lys	Val	Trp	Phe	Ser	Asp	Tyr	Phe	Asn	Ile	Ser	Asp	Thr
		275					280					285			
Ile	Ala	Ile	Ile	Ser	Phe	Phe	Ile	Gly	Phe	Gly	Leu	Arg	Phe	Gly	Ala
	290					295					300				
Lys	Trp	Asn	Phe	Ala	Asn	Ala	Tyr	Asp	Asn	His	Val	Phe	Val	Ala	Gly
305					310					315					320
Arg	Leu	Ile	Tyr	Cys	Leu	Asn	Ile	Ile	Phe	Trp	Tyr	Val	Arg	Leu	Leu
				325					330					335	
Asp	Phe	Leu	Ala	Val	Asn	Gln	Gln	Ala	Gly	Pro	Tyr	Val	Met	Met	Ile
			340					345					350		
Gly	Lys	Met	Val	Ala	Asn	Met	Phe	Tyr	Ile	Val	Val	Ile	Met	Ala	Leu
		355					360					365			
Val	Leu	Leu	Ser	Phe	Gly	Val	Pro	Arg	Lys	Ala	Ile	Leu	Tyr	Pro	His
	370					375					380				
Glu	Ala	Pro	Ser	Trp	Thr	Leu	Ala	Lys	Asp	Ile	Val	Phe	His	Pro	Tyr
385					390					395					400
Trp	Met	Ile	Phe	Gly	Glu	Val	Tyr	Ala	Tyr	Glu	Ile	Asp	Val	Cys	Ala
				405					410					415	
Asn	Asp	Ser	Val	Ile	Pro	Gln	Ile	Cys	Gly	Pro	Gly	Thr	Trp	Leu	Thr
			420					425					430		
Pro	Phe	Leu	Gln	Ala	Val	Tyr	Leu	Phe	Val	Gln	Tyr	Ile	Ile	Met	Val
		435					440					445			
Asn	Leu	Leu	Ile	Ala	Phe	Phe	Asn	Asn	Val	Tyr	Leu	Gln	Val	Lys	Ala
	450					455					460				
Ile	Ser	Asn	Ile	Val	Trp	Lys	Tyr	Gln	Arg	Tyr	His	Phe	Ile	Met	Ala
465					470					475					480
Tyr	His	Glu	Lys	Pro	Val	Leu	Pro	Pro	Pro	Leu	Ile	Ile	Leu	Ser	His
				485					490					495	
Ile	Val	Ser	Leu	Phe	Cys	Cys	Ile	Cys	Lys	Arg	Arg	Lys	Lys	Asp	Lys
			500					505					510		
Thr	Ser	Asp	Gly	Pro	Lys	Leu	Phe	Leu	Thr	Glu	Glu	Asp	Gln	Lys	Lys
		515					520					525			
Leu	His	Asp	Phe	Glu	Glu	Gln	Cys	Val	Glu	Met	Tyr	Phe	Asn	Glu	Lys
	530					535					540				
Asp	Asp	Lys	Phe	His	Ser	Gly	Ser	Glu	Glu	Arg	Ile	Arg	Val	Thr	Phe
545					550					555					560
Glu	Arg	Val	Glu	Gln	Met	Cys	Ile	Gln	Ile	Lys	Glu	Val	Gly	Asp	Arg
				565					570					575	

-35-

Val	Asn	Tyr	Ile	Lys	Arg	Ser	Leu	Gln	Ser	Leu	Asp	Ser	Gln	Ile	Gly
			580					585					590		
His	Leu	Gln	Asp	Leu	Ser	Ala	Leu	Thr	Val	Asp	Thr	Leu	Lys	Thr	Leu
		595					600					605			
Thr	Ala	Gln	Lys	Ala	Ser	Glu	Ala	Ser	Lys	Val	His	Asn	Glu	Ile	Thr
	610					615					620				
Arg	Glu	Leu	Ser	Ile	Ser	Lys	His	Leu	Ala	Gln	Asn	Leu	Ile	Asp	Asp
	625				630					635					640
Gly	Pro	Val	Arg	Pro	Ser	Val	Trp	Lys	Lys	His	Gly	Val	Val	Asn	Thr
				645					650					655	
Leu	Ser	Ser	Ser	Leu	Pro	Gln	Gly	Asp	Leu	Glu	Ser	Asn	Asn	Pro	Phe
				660				665					670		
His	Cys	Asn	Ile	Leu	Met	Lys	Asp	Asp	Lys	Asp	Pro	Gln	Cys	Asn	Ile
		675					680					685			
Phe	Gly	Gln	Asp	Leu	Pro	Ala	Val	Pro	Gln	Arg	Lys	Glu	Phe	Asn	Phe
	690					695				700					
Pro	Glu	Ala	Gly	Ser	Ser	Gly	Ala	Leu	Phe	Pro	Ser	Ala	Val	Ser	
	705				710				715					720	
Pro	Pro	Glu	Leu	Arg	Gln	Arg	Leu	His	Gly	Val	Glu	Leu	Leu	Lys	Ile
				725					730					735	
Phe	Asn	Lys	Asn	Gln	Lys	Leu	Gly	Ser	Ser	Ser	Thr	Ser	Ile	Pro	His
			740					745					750		
Leu	Ser	Ser	Pro	Pro	Thr	Lys	Phe	Phe	Val	Ser	Thr	Pro	Ser	Gln	Pro
		755					760					765			
Ser	Cys	Lys	Ser	His	Leu	Glu	Thr	Gly	Thr	Lys	Asp	Gln	Glu	Thr	Val
	770					775					780				
Cys	Ser	Lys	Ala	Thr	Glu	Gly	Asp	Asn	Xaa	Glu	Phe	Gly	Ala	Phe	Val
	785				790					795					800
Gly	His	Arg	Asp	Ser	Met	Asp	Leu	Gln	Arg	Phe	Lys	Glu	Thr	Ser	Asn
				805					810						815
Lys	Ile	Lys	Ile	Leu	Ser	Asn	Asn	Asn	Thr	Ser	Glu	Asn	Thr	Leu	Lys
			820					825					830		
Arg	Val	Ser	Ser	Leu	Ala	Gly	Phe	Thr	Asp	Cys	His	Arg	Thr	Ser	Ile
		835					840					845			
Pro	Val	His	Ser	Lys	Gln	Ala	Glu	Lys	Ile	Ser	Arg	Arg	Pro	Ser	Thr
		850				855					860				
Glu	Asp	Thr	His	Glu	Val	Asp	Ser	Lys	Ala	Ala	Leu	Ile	Pro	Asp	Trp
	865				870					875					880
Leu	Gln	Asp	Arg	Pro	Ser	Asn	Arg	Glu	Met	Pro	Ser	Glu	Glu	Gly	Thr
				885					890					895	
Leu	Asn	Gly	Leu	Thr	Ser	Pro	Phe	Lys	Pro	Ala	Met	Asp	Thr	Asn	Tyr
				900				905					910		
Tyr	Tyr	Ser	Ala	Val	Glu	Arg	Asn	Asn	Leu	Met	Arg	Leu	Ser	Gln	Ser
		915					920					925			
Ile	Pro	Phe	Thr	Pro	Val	Pro	Pro	Arg	Gly	Glu	Pro	Val	Thr	Val	Tyr
	930					935					940				
Arg	Leu	Glu	Glu	Ser	Ser	Pro	Asn	Ile	Leu	Asn	Asn	Ser	Met	Ser	Ser
	945				950					955					960
Trp	Ser	Gln	Leu	Gly	Leu	Cys	Ala	Lys	Ile	Glu	Phe	Leu	Ser	Lys	Glu
				965					970					975	
Glu	Met	Gly	Gly	Gly	Leu	Arg	Arg	Ala	Val	Lys	Val	Gln	Cys	Thr	Trp
			980					985					990		
Ser	Glu	His	Asp	Ile	Leu	Lys	Ser	Gly	His	Leu	Tyr	Ile	Ile	Lys	Ser
		995					1000					1005			
Phe	Leu	Pro	Glu	Val	Val	Asn	Thr	Trp	Ser	Ser	Ile	Tyr	Lys	Glu	Asp
	1010					1015					1020				
Thr	Val	Leu	His	Leu	Cys	Leu	Arg	Glu	Ile	Gln	Gln	Arg	Ala	Ala	
	1025				1030					1035				104	
Gln	Lys	Leu	Thr	Phe	Ala	Phe	Asn	Gln	Met	Lys	Pro	Lys	Ser	Ile	Pro
				1045					1050					1055	

-36-

Tyr Ser Pro Arg Phe Leu Glu Val Phe Leu Leu Tyr Cys His Ser Ala
 1060 1065 1070
 Gly Gln Trp Phe Ala Val Glu Glu Cys Met Thr Gly Glu Phe Arg Lys
 1075 1080 1085
 Tyr Asn Asn Asn Asn Gly Asp Glu Ile Ile Pro Thr Asn Thr Leu Glu
 1090 1095 1100
 Glu Ile Met Leu Ala Phe Ser His Trp Thr Tyr Glu Tyr Thr Arg Gly
 1105 1110 1115 112
 Glu Leu Leu Val Leu Asp Leu Gln Gly Val Gly Glu Asn Leu Thr Asp
 1125 1130 1135
 Pro Ser Val Ile Lys Ala Glu Glu Lys Arg Ser Cys Asp Met Val Phe
 1140 1145 1150
 Gly Pro Ala Asn Leu Gly Glu Asp Ala Ile Lys Asn Phe Arg Ala Lys
 1155 1160 1165
 His His Cys Asn Ser Cys Cys Arg Lys Leu Lys Leu Pro Asp Leu Lys
 1170 1175 1180
 Arg Asn Asp Tyr Thr Pro Asp Lys Ile Ile Phe Pro Gln Asp Glu Pro
 1185 1190 1195 120
 Ser Asp Leu Asn Leu Gln Pro Gly Asn Ser Thr Lys Glu Ser Glu Ser
 1205 1210 1215
 Ala Asn Ser Val Arg Leu Met Leu
 1220

<210> 25

<211> 2180

<212> DNA

<213> Homo Sapiens

<400> 25

tcgaggccaa	gaattcggca	cgagggcctc	gggcaggccc	cctggagcga	cctgcttctt	60
tgggcactgt	tgctgaacag	ggcacagatg	gccatgtact	tctgggagat	gggttccaat	120
gcagtttcc	agctcttgg	ggcctgtttg	ctgctccggg	tgatggcacg	cctggagcct	180
gacgtcgagg	aggcagcacg	gaggaaagac	ctggcgttca	agtttgaggg	gatgggctt	240
gacctctttg	gcgagtgcta	tcgcagcagt	gaggtgaggg	ctgcccgcct	cctcctccgt	300
cgctgcccgc	tctgggggga	tgccacttgc	ctccagctgg	ccatgcaagc	tgacgcccgt	360
gccttctttg	cccaggatgg	ggtacagtct	ctgctgacac	agaagtgggtg	gggagatatg	420
gccagcacta	caccatctg	ggccctggtt	ctcgccctct	tttgccctcc	actcatctac	480
acccgcctca	tcaccttcag	gaaatcagaa	gaggagccca	cacgggagga	gctagagttt	540
gacatggata	gtgtcattaa	tggggaaggg	cctgtcggga	cggcggaccc	agccgagaag	600
acgcgcgtgg	gggtcccgcg	ccagtcgggc	cgctccgggtt	gctgcggggg	ccgctgcggg	660
gggcgcgggt	gcctacgcgc	ctggttccac	ttctggggcg	cgccggtgac	catcttcatt	720
ggcaacgtgg	tcagctacct	gctgttccctg	ctgcttttct	cgccggtgct	gctcgtggat	780
ttccagccgg	cgccgcccgg	ctccctggag	ctgctgctct	atttctgggc	tttcacgctg	840
ctgtgcgagg	aactgcgcca	gggcctgagc	ggaggcgggg	gcagcctcgc	cagcgggggc	900
cccgggcctg	gccatgcctc	actgagccag	cgctgcgcgc	tctacctcgc	cgacagctgg	960
aaccagtgcg	acctagtggc	tctcacctgc	ttcctcctgg	gcgtgggctg	ccggctgacc	1020
ccgggtttgt	accacctggg	ccgcactgtc	ctctgcatcg	acttcatggt	tttcacggtg	1080
cggtctgttc	acatcttcac	ggtcaacaaa	cagctggggc	ccaagatcgt	catcgtgagc	1140
aagatgatga	aggacgtgtt	cttcttcttc	ttcttctcgc	gcgtgtggct	ggtagcctat	1200
ggcgtggcca	cggaggggct	cctgaggcca	cgggacagtg	acttcccaag	tatcctgcgc	1260
cgctctttct	accgtcccta	cctgcagatc	ttcgggcaga	ttccccagga	ggacatggac	1320
gtggccctca	tgagcacacg	caactgctcg	tcggagcccg	gcttctgggc	acaccctcct	1380
ggggcccagg	cgggcacctg	cgtctcccag	tatgccaaact	ggctgggtgg	gctgctcctc	1440
gtcatcttcc	tgctcgtggc	caacatcctg	ctggtcaact	tgctcattgc	catgttcagt	1500
tacacattcg	gcaaagtaca	gggcaacagc	gatctctact	ggaaggcgca	gcgttaaccg	1560
ctcatccggg	aattccactc	tcggcccgcg	ctggccccgc	cctttatcgt	catctccccc	1620
ttgcgcctcc	tgctcaggca	attgtgcagg	cgaccscgga	gccccagcc	gtcctccccg	1680
gccctcgagc	atttccgggt	ttacctttct	aaggaaagccg	agcggaaagct	gctaacgtgg	1740
gaatcgggtg	ataaggagaa	ctttctgctg	gcacgcgcta	gggacaagcg	ggagagcgac	1800
tccgagmgtc	tgaagcgcac	gtcccagaag	gtggacttgg	cactgaaaca	gctgggacac	1860

-37-

```

atccgcgagt acgaacagcg cctgaaagtg ctggagcggg aggtccagca gtgtacctcg 1920
gcccccgcac ctggtggcct tgtccttgag gtgagcccca tgtccatctg ggccactgtc 1980
aggaccacct ttgggagtgt catccttaca aaccacagca tgcccggctc ctcccagaac 2040
cagtcaccagc ctgggaggat caaggcctgg atcccrggcc gttatccatc tggaggctgc 2100
agggtccttg gggtaacagg gaccacagac ccctcaccac tcacagattc ctcacactgg 2160
ggaaataaag ccatttcaga 2180

```

<210> 26
 <211> 725
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> UNSURE
 <222> (553)...(553)
 <223> UNKNOWN

<221> UNSURE
 <222> (603)...(603)
 <223> UNKNOWN

<400> 26

Ser	Arg	Pro	Arg	Ile	Arg	His	Glu	Gly	Leu	Gly	Gln	Ala	Pro	Trp	Ser	1	5	10	15
Asp	Leu	Leu	Leu	Trp	Ala	Leu	Leu	Leu	Asn	Arg	Ala	Gln	Met	Ala	Met	20	25	30	
Tyr	Phe	Trp	Glu	Met	Gly	Ser	Asn	Ala	Val	Ser	Ser	Ala	Leu	Gly	Ala	35	40	45	
Cys	Leu	Leu	Leu	Arg	Val	Met	Ala	Arg	Leu	Glu	Pro	Asp	Ala	Glu	Glu	50	55	60	
Ala	Ala	Arg	Arg	Lys	Asp	Leu	Ala	Phe	Lys	Phe	Glu	Gly	Met	Gly	Val	65	70	75	80
Asp	Leu	Phe	Gly	Glu	Cys	Tyr	Arg	Ser	Ser	Glu	Val	Arg	Ala	Ala	Arg	85	90	95	
Leu	Leu	Leu	Arg	Arg	Cys	Pro	Leu	Trp	Gly	Asp	Ala	Thr	Cys	Leu	Gln	100	105	110	
Leu	Ala	Met	Gln	Ala	Asp	Ala	Arg	Ala	Phe	Phe	Ala	Gln	Asp	Gly	Val	115	120	125	
Gln	Ser	Leu	Leu	Thr	Gln	Lys	Trp	Trp	Gly	Asp	Met	Ala	Ser	Thr	Thr	130	135	140	
Pro	Ile	Trp	Ala	Leu	Val	Leu	Ala	Phe	Phe	Cys	Pro	Pro	Leu	Ile	Tyr	145	150	155	160
Thr	Arg	Leu	Ile	Thr	Phe	Arg	Lys	Ser	Glu	Glu	Pro	Thr	Arg	Glu		165	170	175	
Glu	Leu	Glu	Phe	Asp	Met	Asp	Ser	Val	Ile	Asn	Gly	Glu	Gly	Pro	Val	180	185	190	
Gly	Thr	Ala	Asp	Pro	Ala	Glu	Lys	Thr	Pro	Leu	Gly	Val	Pro	Arg	Gln	195	200	205	
Ser	Gly	Arg	Pro	Gly	Cys	Cys	Gly	Gly	Arg	Cys	Gly	Gly	Arg	Arg	Cys	210	215	220	
Leu	Arg	Arg	Trp	Phe	His	Phe	Trp	Gly	Ala	Pro	Val	Thr	Ile	Phe	Met	225	230	235	240
Gly	Asn	Val	Val	Ser	Tyr	Leu	Leu	Phe	Leu	Leu	Phe	Ser	Arg	Val		245	250	255	
Leu	Leu	Val	Asp	Phe	Gln	Pro	Ala	Pro	Pro	Gly	Ser	Leu	Glu	Leu	Leu	260	265	270	
Leu	Tyr	Phe	Trp	Ala	Phe	Thr	Leu	Leu	Cys	Glu	Glu	Leu	Arg	Gln	Gly	275	280	285	
Leu	Ser	Gly	Gly	Gly	Gly	Ser	Leu	Ala	Ser	Gly	Gly	Pro	Gly	Pro	Gly	290	295	300	

-38-

His Ala Ser Leu Ser Gln Arg Leu Arg Leu Tyr Leu Ala Asp Ser Trp
 305 310 315 320
 Asn Gln Cys Asp Leu Val Ala Leu Thr Cys Phe Leu Leu Gly Val Gly
 325 330 335
 Cys Arg Leu Thr Pro Gly Leu Tyr His Leu Gly Arg Thr Val Leu Cys
 340 345 350
 Ile Asp Phe Met Val Phe Thr Val Arg Leu Leu His Ile Phe Thr Val
 355 360 365
 Asn Lys Gln Leu Gly Pro Lys Ile Val Ile Val Ser Lys Met Met Lys
 370 375 380
 Asp Val Phe Phe Phe Leu Phe Phe Leu Gly Val Trp Leu Val Ala Tyr
 385 390 395 400
 Gly Val Ala Thr Glu Gly Leu Leu Arg Pro Arg Asp Ser Asp Phe Pro
 405 410 415
 Ser Ile Leu Arg Arg Val Phe Tyr Arg Pro Tyr Leu Gln Ile Phe Gly
 420 425 430
 Gln Ile Pro Gln Glu Asp Met Asp Val Ala Leu Met Glu His Ser Asn
 435 440 445
 Cys Ser Ser Glu Pro Gly Phe Trp Ala His Pro Pro Gly Ala Gln Ala
 450 455 460
 Gly Thr Cys Val Ser Gln Tyr Ala Asn Trp Leu Val Val Leu Leu Leu
 465 470 475 480
 Val Ile Phe Leu Leu Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile
 485 490 495
 Ala Met Phe Ser Tyr Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu
 500 505 510
 Tyr Trp Lys Ala Gln Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg
 515 520 525
 Pro Ala Leu Ala Pro Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu
 530 535 540
 Leu Arg Gln Leu Cys Arg Arg Pro Xaa Ser Pro Gln Pro Ser Ser Pro
 545 550 555 560
 Ala Leu Glu His Phe Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg Lys
 565 570 575
 Leu Leu Thr Trp Glu Ser Val His Lys Glu Asn Phe Leu Leu Ala Arg
 580 585 590
 Ala Arg Asp Lys Arg Glu Ser Asp Ser Glu Xaa Leu Lys Arg Thr Ser
 595 600 605
 Gln Lys Val Asp Leu Ala Leu Lys Gln Leu Gly His Ile Arg Glu Tyr
 610 615 620
 Glu Gln Arg Leu Lys Val Leu Glu Arg Glu Val Gln Gln Cys Thr Ser
 625 630 635 640
 Ala Pro Ala Pro Gly Gly Leu Val Leu Glu Val Ser Pro Met Ser Ile
 645 650 655
 Trp Ala Thr Val Arg Thr Thr Phe Gly Ser Val Ile Leu Thr Asn His
 660 665 670
 Ser Met Pro Gly Ser Ser Gln Asn Gln Ser Gln Pro Gly Arg Ile Lys
 675 680 685
 Ala Trp Ile Pro Gly Arg Tyr Pro Ser Gly Gly Cys Arg Val Leu Gly
 690 695 700
 Val Thr Gly Thr Thr Asp Pro Ser Pro Leu Thr Asp Ser Ser His Trp
 705 710 715 720
 Gly Asn Lys Ala Ile
 725

<210> 27

<211> 7419

<212> DNA

<213> Homo Sapiens

<400> 27

cggggaccga	tccagcctcc	ggactctage	ctaggctttt	gcaaaaagct	atttaggtga	60
cactatagaa	ggtacgcctg	caggtagccg	tccggaattc	ccgggtcgac	ccacgcgtcc	120
gcagccccgt	cgccggcgga	ggcgggcgcg	ggcgcgtncc	ctgtggccag	tcacccggag	180
gagttggtcg	cacaattatg	aaagactcgg	cttctgctgc	tagcgccgga	gctgagttag	240
ttctgagaag	gtttccctgg	gcgttccttg	tccggcgggc	tctgctgccg	cctccggaga	300
cgcttcccga	tagatggcta	caggccgcgg	aggaggagga	ggtggagttg	ctgcccttcc	360
ggagtccgcc	ccgtgaggag	aatgtcccag	aaatcctgga	tagaaagcac	tttgaccaag	420
agggaaatgtg	tatatattat	accaagttcc	aaggaccctc	acagatgcct	tccaggatgt	480
caaattttgtc	agcaactcgt	cagggtgttt	tgtggtcgct	tgggtcaagca	acatgcttgt	540
tttactgcaa	gtcttgccat	gaaatactca	gatgtgaaat	tgggtgacca	ttttaatcag	600
gcaatagaag	aatggtctgt	ggaaaagcat	acagaacaga	gccccacgga	tgcttatgga	660
gtcataaatt	ttcaaggggg	ttctcattcc	tacagagcta	agtatgtgag	gctatcatat	720
gacaccaaac	ctgaagtcct	tctgcaactt	aatggcaaat	ggagttaccc	ggagttaccc	780
aaacttggtta	tctctgtaca	tgggggcatg	cagaaatttg	agcttcaccc	acgaatcaag	840
cagttgcttg	gaaaagggtct	tattaaagct	gcagttacaa	ctggagcctg	gatttttaact	900
ggaggagtaa	acacagggtgt	ggcaaaacat	gttggagatg	ccctcaaaga	acatgcttcc	960
agatcatctc	gaaagatttg	cactatcgga	atgctccat	ggggagtgat	tgaaaacaga	1020
aatgatcttg	ttgggagaga	tgtggttgct	ccttatcaaa	ccttattgaa	ccccctgagc	1080
aaattgaatg	ttttgaataa	tctgcattcc	catttcatat	tgggtgatga	tggcactggt	1140
ggaaagtatg	ggcggaagt	cagactgaga	agagaacttg	aaaaaactat	taatcagcaa	1200
agaattcatg	ctaggattgg	ccagggtgtc	cctgtggtgg	cacttatatt	tgagggtggg	1260
ccaaatggtta	ttctcacagt	tcttgaatac	gggtcagaaa	gccccctgt	tccagtagtt	1320
gtgtgtgaag	gaacaggcag	agctgcagat	ctgctagcgt	atattcataa	acaaacagaa	1380
gaaggaggga	atcttcctga	tgcagcagag	cccgatatta	tttccactat	caaaaaaaca	1440
tttaactttg	gccagaatga	agcatttcat	ttatttcaaa	cactgatgga	gtgcatgaaa	1500
agaaaggagc	ttcttactgt	tttccatatt	agctcagatg	aacatcaaga	tatagatgta	1560
gcaataactta	ctgcaactgt	aaaaggtaact	aatgcactctg	catttgacca	gcttatcctt	1620
acattggcat	gggatagagt	tgacattgcc	aaaaatcatg	tatttgttta	tggacagcag	1680
tggctggttg	gatccttgga	acaagctatg	cctgatgctc	ttgtaattgga	tagagtgtga	1740
tttgtaaaaac	ttcttattga	aaatggagta	cttgatgcata	aattccttac	cattccgaga	1800
ctggaagaac	tttacaacac	taaaagaagt	ccaactaatc	caatgctgtt	tcatcttggt	1860
cgagacgtca	aacagggaaa	tcttctcca	ggatataaga	tcactctgat	tgatatagga	1920
cttgttattg	aatatctcat	gggaggaacc	tacagatgca	cctatactag	gaaacgtttt	1980
cgattaatat	ataatagtct	tgggtgaaat	aatcggaggt	ctggccgaaa	tacctccagc	2040
agcactcctc	agtgcgaaa	gagtcatgaa	tcttttgga	atagggcaga	taaaaaggaa	2100
aaaaatgaggc	ataaccattt	cattaagaca	gcacagccct	tccgaccaa	gattgataca	2160
gttatggaag	aaggaaagaa	gaaaagaacc	aaagatgaaa	ttgtagacat	tgatgatcca	2220
gaaaccaagc	gctttcctta	tccacttaat	gaacttttaa	tttgggcttg	ccttatgaag	2280
aggcagggtca	tggccggttt	tttatggcaa	catggtgaa	aatcaatggc	taaagcatta	2340
gttgccctgta	agatctatcg	ttcaatggca	tatgaagcaa	agcagagtga	cctggtagat	2400
gatacttcag	aagaactaaa	acagtattcc	aatgattttg	gtcagttggc	cgttgaatta	2460
ttagaacagt	ccttcagaca	agatgaaacc	atggctatga	aattgctcac	ttatgaactg	2520
aagaaactgga	gtaattcaac	ctgccttaag	ttagcagttt	cttcaagact	tagacctttt	2580
gtagctcaca	cctgtacaca	aatgttggtta	tctgatattg	ggatgggaag	gctgaatatg	2640
aggaaaaatt	cctggtacaa	ggtcactacta	agcattttag	ttccacctgc	catattgctg	2700
ttagagtata	aaactaaggc	tgaaatgtcc	catatcccac	aatctcaaga	tgctcatcag	2760
atgacaatgg	atgacagcga	aaacaacttt	cagaacataa	cagaagagat	ccccatggaa	2820
gtgtttaaag	aagtacggat	tttggatagt	aatgaaggaa	agaatgagat	ggagatacaa	2880
atgaaatcaa	aaaagcttcc	aattacgcga	aagttttatg	ccttttatca	tgaccaattt	2940
gtaaaaattct	ggtttaaacac	gttggtcatat	ttaggatttc	tgatgcttta	tacattttgtg	3000
gttcttgtag	aaatggaaca	gttaccttca	gttcaagaat	ggattgttat	tgcttatatt	3060
tttacttatg	ccattgagaa	agtccgtgag	atctttatgt	ctgaagctgg	gaaagttaac	3120
cagaagatta	aagtatggtt	tagtgattac	ttcaacatca	gtgatacaat	tgccataaatt	3180
tctttcttca	ttggatttgg	actaagattt	ggagcaaaat	ggaactttgc	aaatgcatat	3240
gataatcatg	ttttgtggc	tggaagatta	atttactgtc	ttaacataat	attttggtat	3300
gtgcgtttgc	tagattttct	agctgtaaat	caacaggcag	gaccttatgt	aatgatgatt	3360
ggaaaaatgg	tggccaatat	gttctacatt	gtagtgtatta	tggctcctgt	attacttagt	3420
tttgggtgtc	ccagaaaggc	aatactttat	cctcatgaag	caccatcttg	gactcttgct	3480
aaagatatag	tttttcaccc	atactggatg	atttttggtg	aagttttatgc	atacgaaatt	3540

gatgtgtgtg	caaagtattc	tggtatccct	caaactctgtg	gtcctgggac	gtgggtgact	3600
ccattttctt	aagcagtcct	cctctttgtg	cagtatatca	ttatggttaa	tcttcttatt	3660
gcatttttca	acaatgtgtg	tttacaagtg	aaggcaattt	ccaatattgt	atggaagtac	3720
cagcgttatc	atcttattat	ggcttatcat	gagaaaccag	ttctgcctcc	tccacttatc	3780
attcttagcc	atatagtttc	tctgttttgc	tgcatatgta	agagaagaaa	gaaagataag	3840
acttccgatg	gacccaaact	tttcttaaca	gaagaagatc	aaaagaaact	tcatgatttt	3900
gaagagcagt	gtgttgaaat	gtatttcaat	gaaaaagatg	acaaatttca	ttctgggagt	3960
gaagagagaa	ttcgtgtcac	ttttgaaaga	gtggaacaga	tgtgcattca	gattaaagaa	4020
gttgagatc	gtgtcaacta	cataaaaaga	tcattacaat	cattagattc	tcaaattggc	4080
catttgcaag	atctttcagc	cctgacggta	gatacattaa	aaacactcac	tgcccagaaa	4140
gcgtcggag	ctagcaaagt	tcataatgaa	atcacacgag	aactgagcat	ttccaaacac	4200
ttggctcaaa	accttattga	tgatggtcct	gtaagacctt	ctgtatggaa	aaagcatggt	4260
gttgtaata	cacttagctc	ctctcttctc	caagggtgat	ttgaaagtaa	taatcctttt	4320
cattgtaata	ttttaatgaa	agatgacaaa	gatccccagt	gtaatatatt	tggtcaagac	4380
ttacctgcag	taccccagag	aaaagaattt	aattttccag	aggctggttc	ctcttctggt	4440
gccttattcc	caagtgtgtg	ttcccctcca	gaactgcgac	agagactaca	tggggtagaa	4500
ctcttaaaaa	tatttaataa	aaatcaaaaa	ttaggcagtt	catctactag	cataccacat	4560
ctgtcatccc	caccaaccaa	attttttggt	agtaacccat	ctcagccaag	ttgcaaaaagc	4620
cacttggaag	ctggaaccaa	agatcaagaa	actgtttgct	ctaaagctac	agaaggagat	4680
aatacagaat	ttggagcatt	tgtaggacac	agagatagca	tggatttaca	gaggtttaaa	4740
gaaacatcaa	acaagataaa	aatactatcc	aataacaata	cttctgaaaa	cactttgaaa	4800
cgagttagtt	ctcttgctgg	atttactgac	tgtcacagaa	cttccattcc	tgttcattca	4860
aaacaagcag	aaaaaatcag	tagaaggcca	tctaccgaag	acactcatga	agtagattcc	4920
aaagcagctt	taataccgga	ttggttacaa	gatagaccat	caaacagaga	aatgccatct	4980
gaagaaggaa	cattaaatgg	tctcacttct	ccatttaagc	cagctatgga	tacaaattac	5040
tattattcag	ctgtggaaag	aaataacttg	atgaggttat	cacagagcat	tccatttaca	5100
cctgtgcctc	caagagggga	gcctgtcaca	gtgtatcggt	tggaaagagag	ttcacccaac	5160
atactaaata	acagcatgtc	ttcttggtca	caactaggcc	tctgtgccaa	aatagagttt	5220
ttaagcaaa	aggagatggg	aggaggttta	cgaagagctg	tcaaagtaca	gtgtacctgg	5280
tcagaacatg	atatcctcaa	atcagggcat	ctttatatta	tcaaactctt	tcttccagag	5340
gtggttaata	catggtcaag	tatttataaa	gaagatcacg	ttctgcatct	ctgtctgaga	5400
gaaattcaac	aacagagagc	agcacaaaag	cttacgtttg	cctttaatca	aatgaaaccc	5460
aaatccatac	catattctcc	aaggttcctt	gaagttttcc	tgctgtattg	ccattcagca	5520
ggacagtggg	ttgctgtgga	agaatgtatg	actggagaat	ttagaaaata	caacaataat	5580
aatggagatg	agattattcc	aactaatact	ctggaagaga	tcatgctagc	ctttagccac	5640
tggacttacg	aatatacaag	aggggagtta	ctggtagctg	atttgcaagg	tgttggtgaa	5700
aatttgactg	acccatctgt	gataaaaagc	gaagaaaaga	gatcctgtga	tatggttttt	5760
ggcccagcaa	atctaggaga	agatgcaatt	aaaaacttca	gagcaaaaca	tactgtaat	5820
tctgtctgta	gaaagcttaa	acttccagat	ctgaagagga	atgattatac	gcctgataaa	5880
attatatttc	aatagctaga	gccttcagat	ttgaatcttc	agcctggaaa	ttccacaaaa	5940
gaatcagaat	caactaattc	tgctcgtctg	atgttataat	attaatatta	ctgaatcatt	6000
ggttttgcc	gcacctcaca	gaaatgttac	tgtgtcactt	ttccctcggg	aggaaattgt	6060
ttggttaatat	agaaagggtg	atgcaagttg	aatttgctga	ctccagcaca	gttaaaagg	6120
caatattctt	ttgacctgat	taatcagtca	gaaagtccct	ataggataga	gctggcagct	6180
gagaaatttt	aaaggtaatt	gataattagt	atttgtaact	ttttaaagg	ctctttgtat	6240
agcagaggat	ctcatattgac	tttgttttga	tgagggtgat	gccctctctt	atgtggtaca	6300
ataccattaa	ccaaaggtag	gtgtccatgc	agattttatt	ggcagctgtt	ttattgccat	6360
tcaactaggg	aaatgaagaa	atcacgcagc	cttttggtta	aatggcagtc	aaaattttcc	6420
tcagtgtatt	tagtgtgttc	agtgtatgata	tactgtgttc	ccaactagat	gcttgttggt	6480
cacgggaagg	gaaatgactt	gttctaattc	taggttcaca	gaggtatgag	aagcctgaac	6540
tgaagaccat	tttcaagagg	gacggtattt	atgaatcagg	gttaggctcc	atatttaaa	6600
atagagccag	tttttttttt	aaatagaacc	caaattgtgt	aaaaatgtta	attgggtttt	6660
ttaaacattg	ttttatcaag	tactgtttaa	gtagaagaaa	gccatggtaa	actgatcat	6720
aacctaaatt	ataaaagcag	aaacctaact	cactcgtcaa	gggaagttac	cttttgagga	6780
aagttaaagt	acttttttcc	ctatctgtat	ctatagcaac	aaccagaaac	ttacaaaact	6840
ctccaaagat	tttattgatt	gttatatcaa	atcagaatgt	aaacatgaac	tcttgcatat	6900
atttaaaatt	gtgttggaac	atttgaacat	gaatgctggt	tgggtactta	agaaatttat	6960
tcagtnggat	tatcattatg	tganactggc	agattgcagt	gcanccttat	gccaataaaa	7020
tgtaatttar	cagcccagaa	tattgttgaa	tattcaacaa	taacaagaaa	agcttttcat	7080
ctaagtttta	tgctttaatt	ttttttcttt	ttttttcttt	ttcttttggt	tccttggtac	7140

-41-

taattttaat	ttttatttgg	aagggagcag	tataaagctt	atttgtattt	agtagtgtat	7200
ctcatagata	cagacaaggc	aagagatgat	aagctgttta	aatagtgktt	aatattgatt	7260
gggggtgggg	agaaagaaaa	agtgtattac	ttaaagatac	tatatacskt	ttktatatca	7320
ttaaattcttt	aaaagaaatn	naataaattt	attgttttnc	aaaaaaaaac	ccnntaaaaa	7380
aaaaagggcg	gccccctctag	aggatccctc	gagggggccc			7419

<210> 28
 <211> 1865
 <212> PRT
 <213> Homo Sapiens

<400> 28

Met	Ser	Gln	Lys	Ser	Trp	Ile	Glu	Ser	Thr	Leu	Thr	Lys	Arg	Glu	Cys
1				5					10					15	
Val	Tyr	Ile	Ile	Pro	Ser	Ser	Lys	Asp	Pro	His	Arg	Cys	Leu	Pro	Gly
			20					25					30		
Cys	Gln	Ile	Cys	Gln	Gln	Leu	Val	Arg	Cys	Phe	Cys	Gly	Arg	Leu	Val
		35				40						45			
Lys	Gln	His	Ala	Cys	Phe	Thr	Ala	Ser	Leu	Ala	Met	Lys	Tyr	Ser	Asp
	50					55					60				
Val	Lys	Leu	Gly	Asp	His	Phe	Asn	Gln	Ala	Ile	Glu	Glu	Trp	Ser	Val
65					70					75				80	
Glu	Lys	His	Thr	Glu	Gln	Ser	Pro	Thr	Asp	Ala	Tyr	Gly	Val	Ile	Asn
			85						90					95	
Phe	Gln	Gly	Gly	Ser	His	Ser	Tyr	Arg	Ala	Lys	Tyr	Val	Arg	Leu	Ser
			100					105					110		
Tyr	Asp	Thr	Lys	Pro	Glu	Val	Ile	Leu	Gln	Leu	Leu	Leu	Lys	Glu	Trp
	115					120						125			
Gln	Met	Glu	Leu	Pro	Lys	Leu	Val	Ile	Ser	Val	His	Gly	Gly	Met	Gln
	130					135					140				
Lys	Phe	Glu	Leu	His	Pro	Arg	Ile	Lys	Gln	Leu	Leu	Gly	Lys	Gly	Leu
145					150					155				160	
Ile	Lys	Ala	Ala	Val	Thr	Thr	Gly	Ala	Trp	Ile	Leu	Thr	Gly	Gly	Val
			165						170					175	
Asn	Thr	Gly	Val	Ala	Lys	His	Val	Gly	Asp	Ala	Leu	Lys	Glu	His	Ala
			180					185					190		
Ser	Arg	Ser	Ser	Arg	Lys	Ile	Cys	Thr	Ile	Gly	Ile	Ala	Pro	Trp	Gly
	195					200						205			
Val	Ile	Glu	Asn	Arg	Asn	Asp	Leu	Val	Gly	Arg	Asp	Val	Val	Ala	Pro
	210				215						220				
Tyr	Gln	Thr	Leu	Leu	Asn	Pro	Leu	Ser	Lys	Leu	Asn	Val	Leu	Asn	Asn
225					230					235				240	
Leu	His	Ser	His	Phe	Ile	Leu	Val	Asp	Asp	Gly	Thr	Val	Gly	Lys	Tyr
			245						250					255	
Gly	Ala	Glu	Val	Arg	Leu	Arg	Arg	Glu	Leu	Glu	Lys	Thr	Ile	Asn	Gln
			260					265					270		
Gln	Arg	Ile	His	Ala	Arg	Ile	Gly	Gln	Gly	Val	Pro	Val	Val	Ala	Leu
	275						280					285			
Ile	Phe	Glu	Gly	Gly	Pro	Asn	Val	Ile	Leu	Thr	Val	Leu	Glu	Tyr	Leu
	290					295					300				
Gln	Glu	Ser	Pro	Pro	Val	Pro	Val	Val	Val	Cys	Glu	Gly	Thr	Gly	Arg
305					310					315				320	
Ala	Ala	Asp	Leu	Leu	Ala	Tyr	Ile	His	Lys	Gln	Thr	Glu	Glu	Gly	Gly
			325						330					335	
Asn	Leu	Pro	Asp	Ala	Ala	Glu	Pro	Asp	Ile	Ile	Ser	Thr	Ile	Lys	Lys
			340					345					350		
Thr	Phe	Asn	Phe	Gly	Gln	Asn	Glu	Ala	Leu	His	Leu	Phe	Gln	Thr	Leu
		355					360					365			

-42-

Met	Glu	Cys	Met	Lys	Arg	Lys	Glu	Leu	Ile	Thr	Val	Phe	His	Ile	Gly
370						375					380				
Ser	Asp	Glu	His	Gln	Asp	Ile	Asp	Val	Ala	Ile	Leu	Thr	Ala	Leu	Leu
385					390					395					400
Lys	Gly	Thr	Asn	Ala	Ser	Ala	Phe	Asp	Gln	Leu	Ile	Leu	Thr	Leu	Ala
			405					410						415	
Trp	Asp	Arg	Val	Asp	Ile	Ala	Lys	Asn	His	Val	Phe	Val	Tyr	Gly	Gln
			420					425					430		
Gln	Trp	Leu	Val	Gly	Ser	Leu	Glu	Gln	Ala	Met	Leu	Asp	Ala	Leu	Val
	435						440					445			
Met	Asp	Arg	Val	Ala	Phe	Val	Lys	Leu	Leu	Ile	Glu	Asn	Gly	Val	Ser
450						455					460				
Met	His	Lys	Phe	Leu	Thr	Ile	Pro	Arg	Leu	Glu	Glu	Leu	Tyr	Asn	Thr
465					470					475					480
Lys	Gln	Gly	Pro	Thr	Asn	Pro	Met	Leu	Phe	His	Leu	Val	Arg	Asp	Val
				485					490					495	
Lys	Gln	Gly	Asn	Leu	Pro	Pro	Gly	Tyr	Lys	Ile	Thr	Leu	Ile	Asp	Ile
			500					505					510		
Gly	Leu	Val	Ile	Glu	Tyr	Leu	Met	Gly	Gly	Thr	Tyr	Arg	Cys	Thr	Tyr
	515						520					525			
Thr	Arg	Lys	Arg	Phe	Arg	Leu	Ile	Tyr	Asn	Ser	Leu	Gly	Gly	Asn	Asn
530						535					540				
Arg	Arg	Ser	Gly	Arg	Asn	Thr	Ser	Ser	Ser	Thr	Pro	Gln	Leu	Arg	Lys
545					550					555					560
Ser	His	Glu	Ser	Phe	Gly	Asn	Arg	Ala	Asp	Lys	Lys	Glu	Lys	Met	Arg
				565					570					575	
His	Asn	His	Phe	Ile	Lys	Thr	Ala	Gln	Pro	Phe	Arg	Pro	Lys	Ile	Asp
			580					585					590		
Thr	Val	Met	Glu	Glu	Gly	Lys	Lys	Arg	Thr	Lys	Asp	Glu	Ile	Val	
	595						600				605				
Asp	Ile	Asp	Asp	Pro	Glu	Thr	Lys	Arg	Phe	Pro	Tyr	Pro	Leu	Asn	Glu
	610					615					620				
Leu	Leu	Ile	Trp	Ala	Cys	Leu	Met	Lys	Arg	Gln	Val	Met	Ala	Arg	Phe
625					630					635					640
Leu	Trp	Gln	His	Gly	Glu	Glu	Ser	Met	Ala	Lys	Ala	Leu	Val	Ala	Cys
				645					650					655	
Lys	Ile	Tyr	Arg	Ser	Met	Ala	Tyr	Glu	Ala	Lys	Gln	Ser	Asp	Leu	Val
			660					665					670		
Asp	Asp	Thr	Ser	Glu	Glu	Leu	Lys	Gln	Tyr	Ser	Asn	Asp	Phe	Gly	Gln
	675						680					685			
Leu	Ala	Val	Glu	Leu	Leu	Glu	Gln	Ser	Phe	Arg	Gln	Asp	Glu	Thr	Met
	690					695					700				
Ala	Met	Lys	Leu	Leu	Thr	Tyr	Glu	Leu	Lys	Asn	Trp	Ser	Asn	Ser	Thr
705					710					715					720
Cys	Leu	Lys	Leu	Ala	Val	Ser	Ser	Arg	Leu	Arg	Pro	Phe	Val	Ala	His
				725					730					735	
Thr	Cys	Thr	Gln	Met	Leu	Leu	Ser	Asp	Met	Trp	Met	Gly	Arg	Leu	Asn
			740					745					750		
Met	Arg	Lys	Asn	Ser	Trp	Tyr	Lys	Val	Ile	Leu	Ser	Ile	Leu	Val	Pro
	755						760					765			
Pro	Ala	Ile	Leu	Leu	Leu	Glu	Tyr	Lys	Thr	Lys	Ala	Glu	Met	Ser	His
	770					775					780				
Ile	Pro	Gln	Ser	Gln	Asp	Ala	His	Gln	Met	Thr	Met	Asp	Asp	Ser	Glu
785					790					795					800
Asn	Asn	Phe	Gln	Asn	Ile	Thr	Glu	Glu	Ile	Pro	Met	Glu	Val	Phe	Lys
				805					810					815	
Glu	Val	Arg	Ile	Leu	Asp	Ser	Asn	Glu	Gly	Lys	Asn	Glu	Met	Glu	Ile
			820					825				830			
Gln	Met	Lys	Ser	Lys	Lys	Leu	Pro	Ile	Thr	Arg	Lys	Phe	Tyr	Ala	Phe
	835						840					845			

-43-

Tyr	His	Ala	Pro	Ile	Val	Lys	Phe	Trp	Phe	Asn	Thr	Leu	Ala	Tyr	Leu
850						855				860					
Gly	Phe	Leu	Met	Leu	Tyr	Thr	Phe	Val	Val	Leu	Val	Gln	Met	Glu	Gln
865					870					875					880
Leu	Pro	Ser	Val	Gln	Glu	Trp	Ile	Val	Ile	Ala	Tyr	Ile	Phe	Thr	Tyr
				885					890					895	
Ala	Ile	Glu	Lys	Val	Arg	Glu	Ile	Phe	Met	Ser	Glu	Ala	Gly	Lys	Val
			900					905					910		
Asn	Gln	Lys	Ile	Lys	Val	Trp	Phe	Ser	Asp	Tyr	Phe	Asn	Ile	Ser	Asp
	915						920					925			
Thr	Ile	Ala	Ile	Ile	Ser	Phe	Phe	Ile	Gly	Phe	Gly	Leu	Arg	Phe	Gly
930					935						940				
Ala	Lys	Trp	Asn	Phe	Ala	Asn	Ala	Tyr	Asp	Asn	His	Val	Phe	Val	Ala
945					950					955					960
Gly	Arg	Leu	Ile	Tyr	Cys	Leu	Asn	Ile	Ile	Phe	Trp	Tyr	Val	Arg	Leu
				965					970					975	
Leu	Asp	Phe	Leu	Ala	Val	Asn	Gln	Gln	Ala	Gly	Pro	Tyr	Val	Met	Met
			980				985						990		
Ile	Gly	Lys	Met	Val	Ala	Asn	Met	Phe	Tyr	Ile	Val	Val	Ile	Met	Ala
	995						1000					1005			
Leu	Val	Leu	Leu	Ser	Phe	Gly	Val	Pro	Arg	Lys	Ala	Ile	Leu	Tyr	Pro
	1010					1015					1020				
His	Glu	Ala	Pro	Ser	Trp	Thr	Leu	Ala	Lys	Asp	Ile	Val	Phe	His	Pro
1025					1030					1035					1040
Tyr	Trp	Met	Ile	Phe	Gly	Glu	Val	Tyr	Ala	Tyr	Glu	Ile	Asp	Val	Cys
				1045					1050					1055	
Ala	Asn	Asp	Ser	Val	Ile	Pro	Gln	Ile	Cys	Gly	Pro	Gly	Thr	Trp	Leu
			1060				1065						1070		
Thr	Pro	Phe	Leu	Gln	Ala	Val	Tyr	Leu	Phe	Val	Gln	Tyr	Ile	Ile	Met
		1075					1080					1085			
Val	Asn	Leu	Leu	Ile	Ala	Phe	Phe	Asn	Asn	Val	Tyr	Leu	Gln	Val	Lys
	1090				1095						1100				
Ala	Ile	Ser	Asn	Ile	Val	Trp	Lys	Tyr	Gln	Arg	Tyr	His	Phe	Ile	Met
1105					1110					1115					1120
Ala	Tyr	His	Glu	Lys	Pro	Val	Leu	Pro	Pro	Pro	Leu	Ile	Ile	Leu	Ser
				1125					1130					1135	
His	Ile	Val	Ser	Leu	Phe	Cys	Cys	Ile	Cys	Lys	Arg	Arg	Lys	Lys	Asp
		1140					1145						1150		
Lys	Thr	Ser	Asp	Gly	Pro	Lys	Leu	Phe	Leu	Thr	Glu	Glu	Asp	Gln	Lys
		1155					1160					1165			
Lys	Leu	His	Asp	Phe	Glu	Glu	Gln	Cys	Val	Glu	Met	Tyr	Phe	Asn	Glu
		1170				1175					1180				
Lys	Asp	Asp	Lys	Phe	His	Ser	Gly	Ser	Glu	Glu	Arg	Ile	Arg	Val	Thr
1185					1190					1195					1200
Phe	Glu	Arg	Val	Glu	Gln	Met	Cys	Ile	Gln	Ile	Lys	Glu	Val	Gly	Asp
				1205					1210					1215	
Arg	Val	Asn	Tyr	Ile	Lys	Arg	Ser	Leu	Gln	Ser	Leu	Asp	Ser	Gln	Ile
			1220				1225						1230		
Gly	His	Leu	Gln	Asp	Leu	Ser	Ala	Leu	Thr	Val	Asp	Thr	Leu	Lys	Thr
		1235					1240					1245			
Leu	Thr	Ala	Gln	Lys	Ala	Ser	Glu	Ala	Ser	Lys	Val	His	Asn	Glu	Ile
		1250				1255						1260			
Thr	Arg	Glu	Leu	Ser	Ile	Ser	Lys	His	Leu	Ala	Gln	Asn	Leu	Ile	Asp
1265					1270					1275					1280
Asp	Gly	Pro	Val	Arg	Pro	Ser	Val	Trp	Lys	Lys	His	Gly	Val	Val	Asn
				1285					1290					1295	
Thr	Leu	Ser	Ser	Ser	Leu	Pro	Gln	Gly	Asp	Leu	Glu	Ser	Asn	Asn	Pro
			1300					1305					1310		
Phe	His	Cys	Asn	Ile	Leu	Met	Lys	Asp	Asp	Lys	Asp	Pro	Gln	Cys	Asn
		1315					1320						1325		

-44-

Ile Phe Gly Gln Asp Leu Pro Ala Val Pro Gln Arg Lys Glu Phe Asn
 1330 1335 1340
 Phe Pro Glu Ala Gly Ser Ser Ser Gly Ala Leu Phe Pro Ser Ala Val
 1345 1350 1355 1360
 Ser Pro Pro Glu Leu Arg Gln Arg Leu His Gly Val Glu Leu Leu Lys
 1365 1370 1375
 Ile Phe Asn Lys Asn Gln Lys Leu Gly Ser Ser Ser Thr Ser Ile Pro
 1380 1385 1390
 His Leu Ser Ser Pro Pro Thr Lys Phe Phe Val Ser Thr Pro Ser Gln
 1395 1400 1405
 Pro Ser Cys Lys Ser His Leu Glu Thr Gly Thr Lys Asp Gln Glu Thr
 1410 1415 1420
 Val Cys Ser Lys Ala Thr Glu Gly Asp Asn Thr Glu Phe Gly Ala Phe
 1425 1430 1435 1440
 Val Gly His Arg Asp Ser Met Asp Leu Gln Arg Phe Lys Glu Thr Ser
 1445 1450 1455
 Asn Lys Ile Lys Ile Leu Ser Asn Asn Thr Ser Glu Asn Thr Leu
 1460 1465 1470
 Lys Arg Val Ser Ser Leu Ala Gly Phe Thr Asp Cys His Arg Thr Ser
 1475 1480 1485
 Ile Pro Val His Ser Lys Gln Ala Glu Lys Ile Ser Arg Arg Pro Ser
 1490 1495 1500
 Thr Glu Asp Thr His Glu Val Asp Ser Lys Ala Ala Leu Ile Pro Asp
 1505 1510 1515 1520
 Trp Leu Gln Asp Arg Pro Ser Asn Arg Glu Met Pro Ser Glu Glu Gly
 1525 1530 1535
 Thr Leu Asn Gly Leu Thr Ser Pro Phe Lys Pro Ala Met Asp Thr Asn
 1540 1545 1550
 Tyr Tyr Tyr Ser Ala Val Glu Arg Asn Asn Leu Met Arg Leu Ser Gln
 1555 1560 1565
 Ser Ile Pro Phe Thr Pro Val Pro Pro Arg Gly Glu Pro Val Thr Val
 1570 1575 1580
 Tyr Arg Leu Glu Glu Ser Ser Pro Asn Ile Leu Asn Asn Ser Met Ser
 1585 1590 1595 1600
 Ser Trp Ser Gln Leu Gly Leu Cys Ala Lys Ile Glu Phe Leu Ser Lys
 1605 1610 1615
 Glu Glu Met Gly Gly Gly Leu Arg Arg Ala Val Lys Val Gln Cys Thr
 1620 1625 1630
 Trp Ser Glu His Asp Ile Leu Lys Ser Gly His Leu Tyr Ile Ile Lys
 1635 1640 1645
 Ser Phe Leu Pro Glu Val Val Asn Thr Trp Ser Ser Ile Tyr Lys Glu
 1650 1655 1660
 Asp Thr Val Leu His Leu Cys Leu Arg Glu Ile Gln Gln Gln Arg Ala
 1665 1670 1675 1680
 Ala Gln Lys Leu Thr Phe Ala Phe Asn Gln Met Lys Pro Lys Ser Ile
 1685 1690 1695
 Pro Tyr Ser Pro Arg Phe Leu Glu Val Phe Leu Leu Tyr Cys His Ser
 1700 1705 1710
 Ala Gly Gln Trp Phe Ala Val Glu Glu Cys Met Thr Gly Glu Phe Arg
 1715 1720 1725
 Lys Tyr Asn Asn Asn Asn Gly Asp Glu Ile Ile Pro Thr Asn Thr Leu
 1730 1735 1740
 Glu Glu Ile Met Leu Ala Phe Ser His Trp Thr Tyr Glu Tyr Thr Arg
 1745 1750 1755 1760
 Gly Glu Leu Leu Val Leu Asp Leu Gln Gly Val Gly Glu Asn Leu Thr
 1765 1770 1775
 Asp Pro Ser Val Ile Lys Ala Glu Glu Lys Arg Ser Cys Asp Met Val
 1780 1785 1790
 Phe Gly Pro Ala Asn Leu Gly Glu Asp Ala Ile Lys Asn Phe Arg Ala
 1795 1800 1805

-45-

Lys His His Cys Asn Ser Cys Cys Arg Lys Leu Lys Leu Pro Asp Leu
 1810 1815 1820
 Lys Arg Asn Asp Tyr Thr Pro Asp Lys Ile Ile Phe Pro Gln Asp Glu
 1825 1830 1835 1840
 Pro Ser Asp Leu Asn Leu Gln Pro Gly Asn Ser Thr Lys Glu Ser Glu
 1845 1850 1855
 Ser Thr Asn Ser Val Arg Leu Met Leu
 1860 1865

<210> 29
 <211> 4061
 <212> DNA
 <213> Homo Sapiens

<400> 29

ggtctggaag	cagagccggc	ggagggagcg	cggggccct	gggctgcagg	aggttgcggc	60
ggccgcggca	gcatggtggt	gccggagaag	gagcagagct	ggatcccca	gatcttcaag	120
aagaagacct	gcacgacgtt	catagttagc	tccacagatc	cgggagggac	cttgtgccag	180
tgtgggcggc	cccggaccgc	ccaccccgca	gtggccatgg	aggatgcctt	cggggcagcc	240
gtggtgaccg	tgtgggacag	cgatgcacac	accacggaga	agcccaccga	tgcctacgga	300
gagctggact	tcacgggggc	cggccgcaag	cacagcaatt	tcctccggct	ctctgaccga	360
acggatccag	ctgcagttaa	tagtctggtc	acacgcacat	ggggcttccg	tgccccgaac	420
ctgggtggtg	cagtgcctggg	gggatcgggg	ggccccgtcc	tccagacctg	gctgcaggac	480
ctgctgcgtc	gtgggctggt	gccccgtgcc	cagagcacag	gagcctggat	tgtcactggg	540
ggtctgcaca	cgggcatcgg	cgggcatggt	ggtgtggctg	tacgggacca	tcagatggcc	600
agcactgggg	gcaccaaggt	ggtggccatg	ggtgtggccc	cctgggggtg	ggtccggaat	660
agagacaccc	tcatacaacc	caagggctcg	ttccctgcga	ggtaccgggtg	gcgcggtgac	720
ccggaggacg	gggtccagtt	tcccctggac	tacaactact	cggccttctt	cctggtggac	780
gacggcacac	acggctgcct	ggggggcgag	aaccgcttcc	gcttgcgcct	ggagtccctac	840
atctcacagc	agaagacggg	cgtgggaggg	actggaattg	acatccctgt	cctgctcctc	900
ctgattgatg	gtgatgagaa	gatgttgacg	cgaatagaga	acgccaccca	ggctcagctc	960
ccatgtctcc	tcgtggctgg	ctcaggggga	gctgcggact	gcctggcgga	gacctgggaa	1020
gacactctgg	ccccaggagg	tgggggagcc	aggcaaggcg	aagcccagaga	tcgaatcagg	1080
cgtttctttc	ccaaagggga	ccttgaggte	ctgcaggccc	aggtggagag	gattatgacc	1140
cggaaaggagc	tcctgacagt	ctattcttct	caggatgggt	ctgaggaatt	cgagaccata	1200
gttttgaagg	cccttgtgaa	ggcctgtggg	agctcggagg	cctcagccta	cctggatgag	1260
ctgcgtttgg	ctgtggcttg	gaaccgcgtg	gacattgccc	agagtgaact	ctttcggggg	1320
gacatccaat	ggcggctcct	ccatctcgaa	gcttccctca	tggacgccct	gctgaatgac	1380
cggcctgagt	tcgtgcgtt	gctcatttcc	gcctgggcca	cttccctgacc	1440	
ccgatgcgcc	tggcccaact	ctacagcgcg	gcgcctcca	actcgtcat	ccgcaacctt	1500
ttggaccagg	cgtccacag	cgcaggcacc	aaagcccag	ccctaaaagg	gggagctgcg	1560
gagctccggc	cccctgacgt	ggggcatgtg	ctgaggatgc	tgctggggaa	gatgtgcgcg	1620
ccgaggtacc	cctccggggg	cgcctgggac	cctcaccag	gccagggctt	cggggagagc	1680
atgtatctgc	tctcggaaca	ggccacctcg	cgcctctcgc	tggatgctgg	cctcgggcag	1740
gccccctgga	gcgacctgct	tctttgggca	ctgttgctga	acagggcaca	gatggccatg	1800
tacttctggg	agatgggttc	caatgcagtt	tcctcagctc	ttggggcctg	tttgcctgctc	1860
cgggtgatgg	cacgcctgga	gcctgacgct	gaggaggcag	cacggaggaa	agacctggcg	1920
ttcaagtttg	aggggatggg	cgttgacctc	tttggcgagt	gctatcgcat	cagtgaagtg	1980
agggctgccc	gcctcctcct	ccgtcgtctg	ccgctctggg	gggatgccac	ttgcctccag	2040
ctggccatgc	aagctgacgc	ccgtgccttc	tttggccagg	atggggtaca	gtctctgctg	2100
acacagaagt	ggtggggaga	tatggccagc	actacaccca	tctgggccct	ggttctcgcc	2160
ttcttttgcc	ctccactcat	ctacaccgc	ctatcacct	tcaggaaatc	agaagaggag	2220
cccacacggg	aggagctaga	gtttgacatg	gatagtgtca	ttaatgggga	agggcctgtc	2280
gggacggcgg	acccagccga	gaagacgccg	ctgggggtcc	cgcgccagtc	gggccgtccg	2340
ggttgctgcg	ggggccgctg	cggggggcgc	cgggtgctac	gccgctgggt	ccacttctgg	2400
ggcgcgccgg	tgaccatctt	catgggcaac	gtggctagct	acctgctgtt	cctgctgctt	2460
ttctcgcggg	tgctgctcgt	ggatttccag	cggcgccgcg	cgggtccctt	ggagctgctg	2520
ctctatttct	gggctttcac	gctgctgtgc	gaggaaactgc	gccagggcct	gagcggaggc	2580
gggggcagcc	tcgccagcgg	gggccccggg	cctggccatg	cctcactgag	ccagcgctg	2640
cgcctctacc	tcgccgacag	ctggaaccag	tgcgacctag	tggctctcac	ctgcttcctc	2700

-46-

```

ctgggcgtgg gctgccggct gaccccgggt ttgtaccacc tgggccgcac tgtcctctgc 2760
atcgacttca tggttttcac ggtgcggctg cttcacatct tcacggtcaa caaacagctg 2820
gggccaaga tcgtcatcgt gagcaagatg atgaaggacg tgttcttctt cctcttcttc 2880
ctcgcggtgt ggctggtagc ctatggcggt gccacggagg ggctcctgag gccacgggac 2940
agtgaacttc caagtatcct gcgccgcgtc ttctaccgtc cctacctgca gatcttcggg 3000
cagattcccc aggaggacat ggacgtggcc ctcattggag acagcaactg ctcgtcggag 3060
cccggcttct gggcacaccc tcctggggcc caggcgggca cctgcgtctc ccagtatgcc 3120
aactggctgg tgggtgctgt cctcgtcatc ttctgctcgt tggccaacat cctgctggtc 3180
aacttgctca ttgccatgtt cagttacaca ttccgcaaaag tacagggcaa cagcgatctc 3240
tactggaagg cgcagcggtt ccgcctcatc cgggaattcc actctcggcc cgcgctggcc 3300
ccgcccttta tcgtcatctc ccacttgctc ctcctgctca ggcaattgtg caggcgaccc 3360
cggagcccc agcgcctctc cccggccctc gagcatttcc gggtttacct ttctaaggaa 3420
gccgagcggg agctgctaac gtgggaatcg gtgcataagg agaactttct gctggcacgc 3480
gctagggaca agcgggagag cgactccgag cgtctgaagc gcacgtccca gaaggtggac 3540
ttggcactga aacagctggg acacatccgc gactacgaac agcgcctgaa agtgctggag 3600
cgggaggtcc agcagtgtag ccgcgtcctg ggggtgggtg ccgagggcct gagccgctct 3660
gccttgctgc ccccaggtgg gccgccaccc cctgacctgc ctgggtccaa agactgagcc 3720
ctgctggcgg acttcaagga gaagccccc caggggattt tgctcctaga gtaaggctca 3780
tctgggcctc ggcccccgca cctggtggcc ttgtccttga ggtgagcccc atgtccatct 3840
gggcaactgt caggaccacc tttgggagtg tcatccttac aaaccacagc atgccggct 3900
cctcccagaa ccagtcaccg cctgggagga tcaaggcctg gatcccgggc cgttatccat 3960
ctggaggctg cagggtcctt ggggtaacag ggaccacaga cccctcacca ctcacagatt 4020
cctcacactg gggaaataaa gccatttcag aggaaaaaaa a 4061

```

<210> 30

<211> 1214

<212> PRT

<213> Homo Sapiens

<400> 30

```

Met Val Val Pro Glu Lys Glu Gln Ser Trp Ile Pro Lys Ile Phe Lys
1      5      10
Lys Lys Thr Cys Thr Thr Phe Ile Val Asp Ser Thr Asp Pro Gly Gly
20      25      30
Thr Leu Cys Gln Cys Gly Arg Pro Arg Thr Ala His Pro Ala Val Ala
35      40      45
Met Glu Asp Ala Phe Gly Ala Ala Val Val Thr Val Trp Asp Ser Asp
50      55      60
Ala His Thr Thr Glu Lys Pro Thr Asp Ala Tyr Gly Glu Leu Asp Phe
65      70      75      80
Thr Gly Ala Gly Arg Lys His Ser Asn Phe Leu Arg Leu Ser Asp Arg
85      90      95
Thr Asp Pro Ala Ala Val Tyr Ser Leu Val Thr Arg Thr Trp Gly Phe
100     105     110
Arg Ala Pro Asn Leu Val Val Ser Val Leu Gly Gly Ser Gly Gly Pro
115     120     125
Val Leu Gln Thr Trp Leu Gln Asp Leu Leu Arg Arg Gly Leu Val Arg
130     135     140
Ala Ala Gln Ser Thr Gly Ala Trp Ile Val Thr Gly Gly Leu His Thr
145     150     155     160
Gly Ile Gly Arg His Val Gly Val Ala Val Arg Asp His Gln Met Ala
165     170     175
Ser Thr Gly Gly Thr Lys Val Val Ala Met Gly Val Ala Pro Trp Gly
180     185     190
Val Val Arg Asn Arg Asp Thr Leu Ile Asn Pro Lys Gly Ser Phe Pro
195     200     205
Ala Arg Tyr Arg Trp Arg Gly Asp Pro Glu Asp Gly Val Gln Phe Pro
210     215     220
Leu Asp Tyr Asn Tyr Ser Ala Phe Phe Leu Val Asp Asp Gly Thr His
225     230     235     240

```

-47-

Gly	Cys	Leu	Gly	Gly	Glu	Asn	Arg	Phe	Arg	Leu	Arg	Leu	Glu	Ser	Tyr		
				245					250					255			
Ile	Ser	Gln	Gln	Lys	Thr	Gly	Val	Gly	Gly	Thr	Gly	Ile	Asp	Ile	Pro		
			260					265					270				
Val	Leu	Leu	Leu	Leu	Ile	Asp	Gly	Asp	Glu	Lys	Met	Leu	Thr	Arg	Ile		
		275					280					285					
Glu	Asn	Ala	Thr	Gln	Ala	Gln	Leu	Pro	Cys	Leu	Leu	Val	Ala	Gly	Ser		
		290				295					300						
Gly	Gly	Ala	Ala	Asp	Cys	Leu	Ala	Glu	Thr	Leu	Glu	Asp	Thr	Leu	Ala		
305					310					315					320		
Pro	Gly	Ser	Gly	Gly	Ala	Arg	Gln	Gly	Glu	Ala	Arg	Asp	Arg	Ile	Arg		
				325				330						335			
Arg	Phe	Phe	Pro	Lys	Gly	Asp	Leu	Glu	Val	Leu	Gln	Ala	Gln	Val	Glu		
			340					345					350				
Arg	Ile	Met	Thr	Arg	Lys	Glu	Leu	Leu	Thr	Val	Tyr	Ser	Ser	Glu	Asp		
		355				360						365					
Gly	Ser	Glu	Glu	Phe	Glu	Thr	Ile	Val	Leu	Lys	Ala	Leu	Val	Lys	Ala		
		370				375					380						
Cys	Gly	Ser	Ser	Glu	Ala	Ser	Ala	Tyr	Leu	Asp	Glu	Leu	Arg	Leu	Ala		
385					390					395					400		
Val	Ala	Trp	Asn	Arg	Val	Asp	Ile	Ala	Gln	Ser	Glu	Leu	Phe	Arg	Gly		
				405					410					415			
Asp	Ile	Gln	Trp	Arg	Ser	Phe	His	Leu	Glu	Ala	Ser	Leu	Met	Asp	Ala		
			420					425					430				
Leu	Leu	Asn	Asp	Arg	Pro	Glu	Phe	Val	Arg	Leu	Leu	Ile	Ser	His	Gly		
		435				440						445					
Leu	Ser	Leu	Gly	His	Phe	Leu	Thr	Pro	Met	Arg	Leu	Ala	Gln	Leu	Tyr		
		450				455					460						
Ser	Ala	Ala	Pro	Ser	Asn	Ser	Leu	Ile	Arg	Asn	Leu	Leu	Asp	Gln	Ala		
465					470					475					480		
Ser	His	Ser	Ala	Gly	Thr	Lys	Ala	Pro	Ala	Leu	Lys	Gly	Gly	Ala	Ala		
				485					490					495			
Glu	Leu	Arg	Pro	Pro	Asp	Val	Gly	His	Val	Leu	Arg	Met	Leu	Leu	Gly		
			500					505					510				
Lys	Met	Cys	Ala	Pro	Arg	Tyr	Pro	Ser	Gly	Gly	Ala	Trp	Asp	Pro	His		
		515				520						525					
Pro	Gly	Gln	Gly	Phe	Gly	Glu	Ser	Met	Tyr	Leu	Leu	Ser	Asp	Lys	Ala		
		530				535					540						
Thr	Ser	Pro	Leu	Ser	Leu	Asp	Ala	Gly	Leu	Gly	Gln	Ala	Pro	Trp	Ser		
545					550					555				560			
Asp	Leu	Leu	Leu	Trp	Ala	Leu	Leu	Leu	Asn	Arg	Ala	Gln	Met	Ala	Met		
				565					570					575			
Tyr	Phe	Trp	Glu	Met	Gly	Ser	Asn	Ala	Val	Ser	Ser	Ala	Leu	Gly	Ala		
			580					585					590				
Cys	Leu	Leu	Leu	Arg	Val	Met	Ala	Arg	Leu	Glu	Pro	Asp	Ala	Glu	Glu		
		595				600						605					
Ala	Ala	Arg	Arg	Lys	Asp	Leu	Ala	Phe	Lys	Phe	Glu	Gly	Met	Gly	Val		
		610				615					620						
Asp	Leu	Phe	Gly	Glu	Cys	Tyr	Arg	Ser	Ser	Glu	Val	Arg	Ala	Ala	Arg		
625				630						635				640			
Leu	Leu	Leu	Arg	Arg	Cys	Pro	Leu	Trp	Gly	Asp	Ala	Thr	Cys	Leu	Gln		
				645					650					655			
Leu	Ala	Met	Gln	Ala	Asp	Ala	Arg	Ala	Phe	Phe	Ala	Gln	Asp	Gly	Val		
			660			665						670					
Gln	Ser	Leu	Leu	Thr	Gln	Lys	Trp	Trp	Gly	Asp	Met	Ala	Ser	Thr	Thr		
		675				680						685					
Pro	Ile	Trp	Ala	Leu	Val	Leu	Ala	Phe	Phe	Cys	Pro	Pro	Leu	Ile	Tyr		
		690				695					700						
Thr	Arg	Leu	Ile	Thr	Phe	Arg	Lys	Ser	Glu	Glu	Glu	Pro	Thr	Arg	Glu		
705					710					715					720		

-48-

Glu	Leu	Glu	Phe	Asp	Met	Asp	Ser	Val	Ile	Asn	Gly	Glu	Gly	Pro	Val	725	730	735
Gly	Thr	Ala	Asp	Pro	Ala	Glu	Lys	Thr	Pro	Leu	Gly	Val	Pro	Arg	Gln	740	745	750
Ser	Gly	Arg	Pro	Gly	Cys	Cys	Gly	Gly	Arg	Cys	Gly	Gly	Arg	Arg	Cys	755	760	765
Leu	Arg	Arg	Trp	Phe	His	Phe	Trp	Gly	Ala	Pro	Val	Thr	Ile	Phe	Met	770	775	780
Gly	Asn	Val	Val	Ser	Tyr	Leu	Leu	Phe	Leu	Leu	Phe	Ser	Arg	Val		785	790	800
Leu	Leu	Val	Asp	Phe	Gln	Pro	Ala	Pro	Pro	Gly	Ser	Leu	Glu	Leu	Leu	805	810	815
Leu	Tyr	Phe	Trp	Ala	Phe	Thr	Leu	Leu	Cys	Glu	Glu	Leu	Arg	Gln	Gly	820	825	830
Leu	Ser	Gly	Gly	Gly	Gly	Ser	Leu	Ala	Ser	Gly	Gly	Pro	Gly	Pro	Gly	835	840	845
His	Ala	Ser	Leu	Ser	Gln	Arg	Leu	Arg	Leu	Tyr	Leu	Ala	Asp	Ser	Trp	850	855	860
Asn	Gln	Cys	Asp	Leu	Val	Ala	Leu	Thr	Cys	Phe	Leu	Leu	Gly	Val	Gly	865	870	875
Cys	Arg	Leu	Thr	Pro	Gly	Leu	Tyr	His	Leu	Gly	Arg	Thr	Val	Leu	Cys	885	890	895
Ile	Asp	Phe	Met	Val	Phe	Thr	Val	Arg	Leu	Leu	His	Ile	Phe	Thr	Val	900	905	910
Asn	Lys	Gln	Leu	Gly	Pro	Lys	Ile	Val	Ile	Val	Ser	Lys	Met	Met	Lys	915	920	925
Asp	Val	Phe	Phe	Phe	Leu	Phe	Phe	Leu	Gly	Val	Trp	Leu	Val	Ala	Tyr	930	935	940
Gly	Val	Ala	Thr	Glu	Gly	Leu	Leu	Arg	Pro	Arg	Asp	Ser	Asp	Phe	Pro	945	950	955
Ser	Ile	Leu	Arg	Arg	Val	Phe	Tyr	Arg	Pro	Tyr	Leu	Gln	Ile	Phe	Gly	965	970	975
Gln	Ile	Pro	Gln	Glu	Asp	Met	Asp	Val	Ala	Leu	Met	Glu	His	Ser	Asn	980	985	990
Cys	Ser	Ser	Glu	Pro	Gly	Phe	Trp	Ala	His	Pro	Pro	Gly	Ala	Gln	Ala	995	1000	1005
Gly	Thr	Cys	Val	Ser	Gln	Tyr	Ala	Asn	Trp	Leu	Val	Val	Leu	Leu	Leu	1010	1015	1020
Val	Ile	Phe	Leu	Leu	Val	Ala	Asn	Ile	Leu	Leu	Val	Asn	Leu	Leu	Ile	1025	1030	1035
Ala	Met	Phe	Ser	Tyr	Thr	Phe	Gly	Lys	Val	Gln	Gly	Asn	Ser	Asp	Leu	1045	1050	1055
Tyr	Trp	Lys	Ala	Gln	Arg	Tyr	Arg	Leu	Ile	Arg	Glu	Phe	His	Ser	Arg	1060	1065	1070
Pro	Ala	Leu	Ala	Pro	Pro	Phe	Ile	Val	Ile	Ser	His	Leu	Arg	Leu	Leu	1075	1080	1085
Leu	Arg	Gln	Leu	Cys	Arg	Arg	Pro	Arg	Ser	Pro	Gln	Pro	Ser	Ser	Pro	1090	1095	1100
Ala	Leu	Glu	His	Phe	Arg	Val	Tyr	Leu	Ser	Lys	Glu	Ala	Glu	Arg	Lys	1105	1110	1115
Leu	Leu	Thr	Trp	Glu	Ser	Val	His	Lys	Glu	Asn	Phe	Leu	Leu	Ala	Arg	1125	1130	1135
Ala	Arg	Asp	Lys	Arg	Glu	Ser	Asp	Ser	Glu	Arg	Leu	Lys	Arg	Thr	Ser	1140	1145	1150
Gln	Lys	Val	Asp	Leu	Ala	Leu	Lys	Gln	Leu	Gly	His	Ile	Arg	Glu	Tyr	1155	1160	1165
Glu	Gln	Arg	Leu	Lys	Val	Leu	Glu	Arg	Glu	Val	Gln	Gln	Cys	Ser	Arg	1170	1175	1180
Val	Leu	Gly	Trp	Val	Ala	Glu	Ala	Leu	Ser	Arg	Ser	Ala	Leu	Leu	Pro	1185	1190	1195
																		1200

Pro Gly Gly Pro Pro Pro Pro Asp Leu Pro Gly Ser Lys Asp
1205 1210

<210> 31
<211> 4646
<212> DNA
<213> Homo Sapiens

<400> 31
tcgacccacg cgtccgcccc cgcgtccgcc caccgcgtccg cccacgcgtc cgccccacgcg 60
tccgccccacg cgtccgggggt gaaagmramy cmgygktsms aaaaaccgtc acttaggaaa 120
agatgtcctt tcgggcagcc aggctcagca tgaggaaacag aaggaaatgac actctggaca 180
gcacccggac cctgtactcc agcgcgtctc ggagcacaga cttgtcttac agtgaaagcg 240
acttggtgaa ttttattcaa gcaaatttta agaaacgaga atgtgtcttc ttaccaaaag 300
attccaaaggc caccgagaat gtgtgcaagt gtggctatgc ccagagccag cacatggaag 360
gcacccagat caaccaaagt gagaaatgga actacaagaa acacaccaag gaatttccta 420
ccgacgcctt tggggatatt cagtttgaga cactggggaa gaaagggag tatatacgtc 480
tgtctgcga caccgacgcg gaaatccttt acgagctgct gaccacgac tggcacctga 540
aaacacccaa cctggtcatt tctgtgaccg ggggcgcca gaacttcgcc ctgaagccgc 600
gcatgcgcaa gatcttcagc cggctcatct acatcgcgca gtccaaagg gcttggattc 660
tcacgggagg caccattat ggcctgatga agtacatcg ggaggtgggt agagataaca 720
ccatcagcag gagttcagag gagaatattg tggccattgg catagcagct tggggcatgg 780
tctccaaccg ggacaccctc atcaggaatt gcgagtctga gggctatatt ttagccaggt 840
accttatgga tgacttcaca agagatccac tgtgtatcct ggacaacaac cacacacatt 900
tgctgctcgt ggacaatggc tgtcatggac atcccactgt cgaagcaaa ctccgggaatc 960
agctagagaa gtatatctct gagcgacta ttcaagatt caactatggt ggcaagatcc 1020
ccattgtgtg ttttgcccaa ggaggtggaa aagagacttt gaaagccatc aatacctcca 1080
tcaaaaataa aattccttgt gtgggtgggt aaggctcggg ccagatcgct gatgtgatcg 1140
ctagcctggt ggaggtggag gatgccctga catcttctgc cgtcaaggag aagctgggtc 1200
gctttttacc ccgcacggtg tcccggtctc ctgaggagga gactgagagt tggatcaaat 1260
ggctcaaaga aattctcgaa tgttctcacc tattaacagt tattaataatg gaagaactg 1320
gggatgaaat tgtgagcaat gccatctcct acgctctata caaagccttc agcaccagt 1380
agcaagacaa ggataactgg aatgggcagc tgaagcttct gctggagtgg aaccagctgg 1440
acttagccaa tgatgagatt ttcaccaatg accgccgatg ggagtctgct gaccttcaag 1500
aagtcattgt tacggctctc ataaaggaca gaccaaagt tgtccgcctc ttctcgaga 1560
atggcttgaa cctacggaa gttctcacc atgatgtcct cactgaactc ttctccaacc 1620
acttcagcac gcttgtgtac cggaaatctgc agatcgccaa gaattcctat aatgatgcc 1680
tcctcacgtt tgtctggaat ctggttgcca acttccgaag aggttcccg aaggaagaca 1740
gaaatggccg ggacgagatg gacatagaac tccacgacgt gtctcctatt actcgccacc 1800
ccctgcaagc cctcttcac tgggccattc tcagaataa gaaggaactc tccaaagtca 1860
tttgggagca gaccaggggc tgcactctgg cagccctggg agccagcaag cttctgaaga 1920
ctctggccaa agtgaagaac gacatcaatg ctgctgggga gtccgaggag ctggctaatt 1980
agtacgagac ccgggctgtt gagctgttca ctgagtgtta cagcagcgat gaagacttg 2040
cagaacagct gctggtctat tctgtgaa gcttgggggt aagcaactgt ctggagctgg 2100
cgggtggagc cacagaccag catttcatcg cccagcctgg ggtccagaat tttctttcta 2160
agcaatggta tggagagatt tcccagagaca ccaagaactg gaagattatc ctgtgtctgt 2220
ttattatacc cttggtgggc tgtggctttg tatcatttag gaagaaacct gtcgacaagc 2280
acaagaagct gctttggtac tatgtggcgt tcttcacctc ccccttcgtg gtcttctcct 2340
ggaatgtggt cttctacac gccttctctc tctgtttgc ctacgtgctg ctcatggatt 2400
tccattcggg gccacacccc cccgagctgg tctgtactc gctggctctt gtctcttct 2460
gtgatgaagt gagacagtgg tacgtaaatg ggtgaaatta ttttactgac ctgtggaatg 2520
tgatggacac gctgggctt ttttacttca tagcaggaat tgtatttcgg ctccactctt 2580
ctaataaaaag ctctttgtat tctggacgag tcattttctg tctggactac attattttca 2640
ctctaagatt gatccacatt tttactgtaa gcagaaactt aggaccaag attataatgc 2700
tgacagaggat gctgatcgat gtgttcttct tctgttctct ctttgcggtg tggatgggtg 2760
cctttggcgt ggccaggcaa gggatcctta ggcagaatga gcagcgctgg aggtggatat 2820
tccgttcggc catctacgag cctacctgg ccagtgtcgg ccagtgtcgg agtgacgtgg 2880
atggtaccac gtatgacttt gccactgca ccttcaactg gaatgagtc aagccactgt 2940
gtgtggagct ggatgagcac aacctgcccc ggttccccga gtggatcacc atccccctgg 3000
tgtgcatcta catgttatcc accaacatcc tgctggtcaa cctgctggtc gccatgtttg 3060

-50-

gctacacggt	gggcaccgctc	caggagaaca	atgaccaggt	ctggaagttc	cagaggtact	3120
tcctgggtgca	ggagtactgc	agccgcctca	atatccccctt	ccccttcac	gtcttcgctt	3180
acttctacat	ggtgggtgaag	aagtgccttc	agtgttgctg	caaggagaaa	aacatggagt	3240
cttctgtctg	ctgtttcaaa	aatgaagaca	atgagactct	ggcatgggag	ggtgtcatga	3300
aggaaaacta	ccttgtcaag	atcaacacaa	aagccaacga	cacctcagag	gaaatgagggc	3360
atcgatttag	acaactggat	acaaagctta	atgatctcaa	gggtcttctg	aaagagattg	3420
ctaataaaat	caaataaaac	tgtatgaact	ctaattggaga	aaaatcta	tatagcaaga	3480
tcatattaag	gaatgctgat	gaacaatttt	gctatcgact	actaatgag	agattttcag	3540
accctgggt	acatgggtgga	tgattttaaa	tcaccctagt	gtgctgagac	cttgagaata	3600
aagtgtgtga	ttggtttcat	acttgaagac	ggatataaag	gaagaatatt	tcctttatgt	3660
gtttctccag	aatggtgcct	gtttctctct	gtgtctcaat	gcctgggact	ggaggttgat	3720
agtttaagt	tggtcttacc	gcctcctttt	tcctttaatc	ttatttttga	tgaacacata	3780
tataggagaa	catctatcct	atgaataaga	acctgggtcat	gctttactcc	tgtattgtta	3840
ttttgttcat	ttccaattga	ttctctactt	ttcccttttt	tgtattatgt	gactaattag	3900
ttggcatatt	gtwaaaagtc	tctcaaatta	ggccagattc	taaaacatgc	tgcagcaaga	3960
ggaccccgct	ctcttcagga	aaagtgtttt	catttctcag	gatgcttctt	acctgtcaga	4020
ggaggtgaca	aggcagtcct	ttgctctctt	ggactcacca	ggctcctatt	gaaggaaacca	4080
ccccattcc	taaatatgtg	aaaagtcgcc	aaaaatgcaa	ccttgaaagg	cactactgac	4140
tttgttctta	ttggatactc	ctcttattta	ttatttttcc	attaaaaata	atagctggct	4200
attatagaaa	atttagacca	tacagagatg	tagaaagaac	ataaattgtc	cccattacct	4260
taaggtaatc	actgctaaca	atttctggat	ggtttttcaa	gtctattttt	tttctatgta	4320
tgctctcaatt	ctctttcaaa	tttttacaga	atgttatcat	actacatata	tactttttat	4380
gtaagctttt	tcacttagta	ttttatcaaa	tatgttttta	ttatattcat	agccttctta	4440
aacatttatat	caataattgc	ataataggca	acctctagcg	attaccataa	ttttgctcat	4500
tgaaggetat	ctccagttga	tcattgggat	gagcatcttt	gtgcatgaat	cctattgctg	4560
tatttgggaa	aattttccaa	ggttagattc	caataaatat	ctatttatta	ttaaaaaa	4620
aaaaaaaagg	gcggccgctc	tagagt				4646

<210> 32

<211> 1104

<212> PRT

<213> Homo Sapiens

<400> 32

Met	Ser	Phe	Arg	Ala	Ala	Arg	Leu	Ser	Met	Arg	Asn	Arg	Arg	Asn	Asp
1				5					10					15	
Thr	Leu	Asp	Ser	Thr	Arg	Thr	Leu	Tyr	Ser	Ser	Ala	Ser	Arg	Ser	Thr
			20					25					30		
Asp	Leu	Ser	Tyr	Ser	Glu	Ser	Asp	Leu	Val	Asn	Phe	Ile	Gln	Ala	Asn
		35					40					45			
Phe	Lys	Lys	Arg	Glu	Cys	Val	Phe	Phe	Thr	Lys	Asp	Ser	Lys	Ala	Thr
	50					55					60				
Glu	Asn	Val	Cys	Lys	Cys	Gly	Tyr	Ala	Gln	Ser	Gln	His	Met	Glu	Gly
	65					70				75				80	
Thr	Gln	Ile	Asn	Gln	Ser	Glu	Lys	Trp	Asn	Tyr	Lys	Lys	His	Thr	Lys
				85					90					95	
Glu	Phe	Pro	Thr	Asp	Ala	Phe	Gly	Asp	Ile	Gln	Phe	Glu	Thr	Leu	Gly
			100					105					110		
Lys	Lys	Gly	Lys	Tyr	Ile	Arg	Leu	Ser	Cys	Asp	Thr	Asp	Ala	Glu	Ile
		115					120					125			
Leu	Tyr	Glu	Leu	Leu	Thr	Gln	His	Trp	His	Leu	Lys	Thr	Pro	Asn	Leu
	130					135						140			
Val	Ile	Ser	Val	Thr	Gly	Gly	Ala	Lys	Asn	Phe	Ala	Leu	Lys	Pro	Arg
	145				150					155				160	
Met	Arg	Lys	Ile	Phe	Ser	Arg	Leu	Ile	Tyr	Ile	Ala	Gln	Ser	Lys	Gly
				165					170					175	
Ala	Trp	Ile	Leu	Thr	Gly	Gly	Thr	His	Tyr	Gly	Leu	Met	Lys	Tyr	Ile
			180					185					190		
Gly	Glu	Val	Val	Arg	Asp	Asn	Thr	Ile	Ser	Arg	Ser	Ser	Glu	Glu	Asn
		195					200					205			

-51-

Ile	Val	Ala	Ile	Gly	Ile	Ala	Ala	Trp	Gly	Met	Val	Ser	Asn	Arg	Asp
210						215					220				
Thr	Leu	Ile	Arg	Asn	Cys	Asp	Ala	Glu	Gly	Tyr	Phe	Leu	Ala	Gln	Tyr
225					230					235					240
Leu	Met	Asp	Asp	Phe	Thr	Arg	Asp	Pro	Leu	Cys	Ile	Leu	Asp	Asn	Asn
				245					250					255	
His	Thr	His	Leu	Leu	Leu	Val	Asp	Asn	Gly	Cys	His	Gly	His	Pro	Thr
			260					265					270		
Val	Glu	Ala	Lys	Leu	Arg	Asn	Gln	Leu	Glu	Lys	Tyr	Ile	Ser	Glu	Arg
		275					280					285			
Thr	Ile	Gln	Asp	Ser	Asn	Tyr	Gly	Gly	Lys	Ile	Pro	Ile	Val	Cys	Phe
290					295						300				
Ala	Gln	Gly	Gly	Gly	Lys	Glu	Thr	Leu	Lys	Ala	Ile	Asn	Thr	Ser	Ile
305					310					315					320
Lys	Asn	Lys	Ile	Pro	Cys	Val	Val	Val	Glu	Gly	Ser	Gly	Gln	Ile	Ala
				325					330					335	
Asp	Val	Ile	Ala	Ser	Leu	Val	Glu	Val	Glu	Asp	Ala	Leu	Thr	Ser	Ser
			340					345					350		
Ala	Val	Lys	Glu	Lys	Leu	Val	Arg	Phe	Leu	Pro	Arg	Thr	Val	Ser	Arg
		355					360					365			
Leu	Pro	Glu	Glu	Glu	Thr	Glu	Ser	Trp	Ile	Lys	Trp	Leu	Lys	Glu	Ile
370					375						380				
Leu	Glu	Cys	Ser	His	Leu	Leu	Thr	Val	Ile	Lys	Met	Glu	Glu	Ala	Gly
385					390					395					400
Asp	Glu	Ile	Val	Ser	Asn	Ala	Ile	Ser	Tyr	Ala	Leu	Tyr	Lys	Ala	Phe
				405					410					415	
Ser	Thr	Ser	Glu	Gln	Asp	Lys	Asp	Asn	Trp	Asn	Gly	Gln	Leu	Lys	Leu
			420					425					430		
Leu	Leu	Glu	Trp	Asn	Gln	Leu	Asp	Leu	Ala	Asn	Asp	Glu	Ile	Phe	Thr
		435					440					445			
Asn	Asp	Arg	Arg	Trp	Glu	Ser	Ala	Asp	Leu	Gln	Glu	Val	Met	Phe	Thr
450					455						460				
Ala	Leu	Ile	Lys	Asp	Arg	Pro	Lys	Phe	Val	Arg	Leu	Phe	Leu	Glu	Asn
465					470					475					480
Gly	Leu	Asn	Leu	Arg	Lys	Phe	Leu	Thr	His	Asp	Val	Leu	Thr	Glu	Leu
				485				490						495	
Phe	Ser	Asn	His	Phe	Ser	Thr	Leu	Val	Tyr	Arg	Asn	Leu	Gln	Ile	Ala
			500					505					510		
Lys	Asn	Ser	Tyr	Asn	Asp	Ala	Leu	Thr	Phe	Val	Trp	Lys	Leu	Val	
		515					520				525				
Ala	Asn	Phe	Arg	Arg	Gly	Phe	Arg	Lys	Glu	Asp	Arg	Asn	Gly	Arg	Asp
		530			535					540					
Glu	Met	Asp	Ile	Glu	Leu	His	Asp	Val	Ser	Pro	Ile	Thr	Arg	His	Pro
545					550					555					560
Leu	Gln	Ala	Leu	Phe	Ile	Trp	Ala	Ile	Leu	Gln	Asn	Lys	Lys	Glu	Leu
			565					570						575	
Ser	Lys	Val	Ile	Trp	Glu	Gln	Thr	Arg	Gly	Cys	Thr	Leu	Ala	Ala	Leu
			580					585					590		
Gly	Ala	Ser	Lys	Leu	Leu	Lys	Thr	Leu	Ala	Lys	Val	Lys	Asn	Asp	Ile
		595					600					605			
Asn	Ala	Ala	Gly	Glu	Ser	Glu	Glu	Leu	Ala	Asn	Glu	Tyr	Glu	Thr	Arg
		610				615					620				
Ala	Val	Glu	Leu	Phe	Thr	Glu	Cys	Tyr	Ser	Ser	Asp	Glu	Asp	Leu	Ala
625					630					635					640
Glu	Gln	Leu	Leu	Val	Tyr	Ser	Cys	Glu	Ala	Trp	Gly	Gly	Ser	Asn	Cys
			645					650						655	
Leu	Glu	Leu	Ala	Val	Glu	Ala	Thr	Asp	Gln	His	Phe	Ile	Ala	Gln	Pro
		660						665					670		
Gly	Val	Gln	Asn	Phe	Leu	Ser	Lys	Gln	Trp	Tyr	Gly	Glu	Ile	Ser	Arg
		675					680					685			

-52-

Asp	Thr	Lys	Asn	Trp	Lys	Ile	Ile	Leu	Cys	Leu	Phe	Ile	Ile	Pro	Leu
690						695					700				
Val	Gly	Cys	Gly	Phe	Val	Ser	Phe	Arg	Lys	Lys	Pro	Val	Asp	Lys	His
705					710					715					720
Lys	Lys	Leu	Leu	Trp	Tyr	Tyr	Val	Ala	Phe	Phe	Thr	Ser	Pro	Phe	Val
				725					730					735	
Val	Phe	Ser	Trp	Asn	Val	Val	Phe	Tyr	Ile	Ala	Phe	Leu	Leu	Leu	Phe
			740					745					750		
Ala	Tyr	Val	Leu	Leu	Met	Asp	Phe	His	Ser	Val	Pro	His	Pro	Pro	Glu
		755					760					765			
Leu	Val	Leu	Tyr	Ser	Leu	Val	Phe	Val	Leu	Phe	Cys	Asp	Glu	Val	Arg
		770				775					780				
Gln	Trp	Tyr	Val	Asn	Gly	Val	Asn	Tyr	Phe	Thr	Asp	Leu	Trp	Asn	Val
785					790						795				800
Met	Asp	Thr	Leu	Gly	Leu	Phe	Tyr	Phe	Ile	Ala	Gly	Ile	Val	Phe	Arg
				805					810					815	
Leu	His	Ser	Ser	Asn	Lys	Ser	Ser	Leu	Tyr	Ser	Gly	Arg	Val	Ile	Phe
			820					825					830		
Cys	Leu	Asp	Tyr	Ile	Ile	Phe	Thr	Leu	Arg	Leu	Ile	His	Ile	Phe	Thr
		835					840					845			
Val	Ser	Arg	Asn	Leu	Gly	Pro	Lys	Ile	Ile	Met	Leu	Gln	Arg	Met	Leu
		850				855					860				
Ile	Asp	Val	Phe	Phe	Phe	Leu	Phe	Leu	Phe	Ala	Val	Trp	Met	Val	Ala
865					870					875					880
Phe	Gly	Val	Ala	Arg	Gln	Gly	Ile	Leu	Arg	Gln	Asn	Glu	Gln	Arg	Trp
				885					890					895	
Arg	Trp	Ile	Phe	Arg	Ser	Val	Ile	Tyr	Glu	Pro	Tyr	Leu	Ala	Met	Phe
			900					905					910		
Gly	Gln	Val	Pro	Ser	Asp	Val	Asp	Gly	Thr	Thr	Tyr	Asp	Phe	Ala	His
		915					920					925			
Cys	Thr	Phe	Thr	Gly	Asn	Glu	Ser	Lys	Pro	Leu	Cys	Val	Glu	Leu	Asp
		930				935					940				
Glu	His	Asn	Leu	Pro	Arg	Phe	Pro	Glu	Trp	Ile	Thr	Ile	Pro	Leu	Val
945					950					955					960
Cys	Ile	Tyr	Met	Leu	Ser	Thr	Asn	Ile	Leu	Leu	Val	Asn	Leu	Leu	Val
				965					970					975	
Ala	Met	Phe	Gly	Tyr	Thr	Val	Gly	Thr	Val	Gln	Glu	Asn	Asn	Asp	Gln
			980					985					990		
Val	Trp	Lys	Phe	Gln	Arg	Tyr	Phe	Leu	Val	Gln	Glu	Tyr	Cys	Ser	Arg
		995					1000					1005			
Leu	Asn	Ile	Pro	Phe	Pro	Phe	Ile	Val	Phe	Ala	Tyr	Phe	Tyr	Met	Val
		1010				1015					1020				
Val	Lys	Lys	Cys	Phe	Lys	Cys	Cys	Cys	Lys	Glu	Lys	Asn	Met	Glu	Ser
1025					1030					1035					104
Ser	Val	Cys	Cys	Phe	Lys	Asn	Glu	Asp	Asn	Glu	Thr	Leu	Ala	Trp	Glu
				1045					1050					1055	
Gly	Val	Met	Lys	Glu	Asn	Tyr	Leu	Val	Lys	Ile	Asn	Thr	Lys	Ala	Asn
			1060					1065					1070		
Asp	Thr	Ser	Glu	Glu	Met	Arg	His	Arg	Phe	Arg	Gln	Leu	Asp	Thr	Lys
		1075					1080					1085			
Leu	Asn	Asp	Leu	Lys	Gly	Leu	Leu	Lys	Glu	Ile	Ala	Asn	Lys	Ile	Lys
		1090					1095					1100			

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau



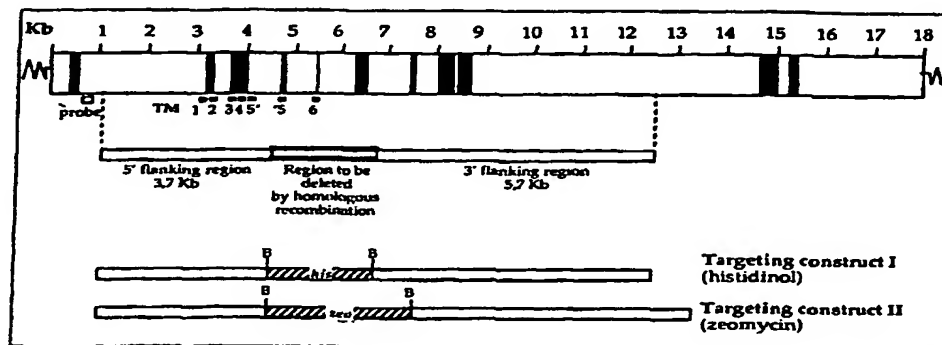
(43) International Publication Date
13 July 2000 (13.07.2000)

PCT

(10) International Publication Number
WO 00/40614 A3

- (51) International Patent Classification⁷: C07K 14/705, C12N 15/12, C12Q 1/68, C12N 5/10, C07K 16/28, G01N 33/53, A61K 38/17
- (74) Agent: PLUMER, Elizabeth, R.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).
- (21) International Application Number: PCT/US99/29996
- (81) Designated States (*national*): AU, CA, JP, US.
- (22) International Filing Date:
20 December 1999 (20.12.1999)
- (84) Designated States (*regional*): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).
- (25) Filing Language: English
- Published:
— with international search report
- (26) Publication Language: English
- (30) Priority Data:
60/114,220 30 December 1998 (30.12.1998) US
60/120,018 29 January 1999 (29.01.1999) US
60/140,415 22 June 1999 (22.06.1999) US
- (88) Date of publication of the international search report:
22 February 2001
- (48) Date of publication of this corrected version:
30 August 2001
- (71) Applicant (*for all designated States except US*): BETH ISRAEL DEACONESS MEDICAL CENTER, INC. [US/US]; 1 Deaconess Road, Boston, MA 02215 (US).
- (15) Information about Correction:
see PCT Gazette No. 35/2001 of 30 August 2001, Section II
- (72) Inventor; and
- (75) Inventor/Applicant (*for US only*): SCHARENBERG, Andrew, M. [US/US]; 12 Skyview Road, Lexington, MA 02420 (US).
- For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

(54) Title: CHARACTERIZATION OF THE SOC/CRAC CALCIUM CHANNEL PROTEIN FAMILY



(57) Abstract: Nucleic acids encoding SOC/CRAC calcium channel polypeptides, including fragments and biologically functional variants thereof and encoded polypeptides are provided. The nucleic acids and polypeptides disclosed herein are useful as therapeutic and diagnostic agents. Agents that selectively bind to the foregoing polypeptides and genes also are provided.

CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY**Field of the Invention**

This invention relates to nucleic acids coding for a novel family of calcium channel polypeptides, the encoded polypeptides, unique fragments of the foregoing, and methods of making and using same.

Background of the Invention

Calcium channels are membrane-spanning, multi-subunit proteins that facilitate the controlled transport ("flux") of Ca^{2+} ions into and out of cells. Cells throughout the animal kingdom, and at least some bacterial, fungal and plant cells, possess one or more types of calcium channels. In general, "excitable" cells, such as neurons of the central nervous system, peripheral nerve cells, and muscle cells, including those of skeletal muscles, cardiac muscles, and venous and arterial smooth muscles, possess voltage-dependent calcium channels. In a voltage-dependent calcium channel, the transport of Ca^{2+} ions into and out of the cells requires a certain minimal level of depolarization (the difference in potential between the inside of the cell bearing the channel and the extracellular environment) with the rate of Ca^{2+} cell flux dependent on the difference in potential. In "non-excitable" cells, calcium influx is thought to occur predominantly in response to stimuli which cause the release of calcium from intracellular stores. This process, termed *store operated calcium influx*, is not well understood.

Characterization of a particular type of calcium channel by analysis of whole cells is complicated by the presence of mixed populations of different types of calcium channels in the majority of cells. Although single-channel recording methods can be used to examine individual calcium channels, such analysis does not reveal information related to the molecular structure or biochemical composition of the channel. Furthermore, in this type of analysis, the channel is isolated from other cellular constituents that might be important for the channel's natural functions and pharmacological interactions. To study the calcium channel structure-function relationship, large amounts of pure channel protein are needed. However, acquiring large amounts of pure protein is difficult in view of the complex nature of these multisubunit proteins, the varying concentrations of calcium channel proteins in tissue sources, the presence of mixed populations of calcium channel proteins in tissues, and the modifications of the native protein that can occur during the isolation procedure.

Summary of the Invention

The invention is based on the identification of a novel family of calcium channel polypeptides and the molecular cloning and partial characterization of a novel member of this family that is expressed predominantly in human hematopoietic cells, liver, and kidney. This newly identified family of calcium channel polypeptides is designated, "SOC" or "CRAC" or "ICRAC", for Store Operated Channels or Calcium Release Activated Channels. Although not wishing to be bound to any particular theory or mechanism, it is believed that the SOC/CRAC calcium channel polypeptides are transmembrane polypeptides that modulate Ca^{2+} flux "into" and "out of" a cell, for example, in certain instances they may be activated upon depletion of Ca^{2+} from intracellular calcium stores, allowing Ca^{2+} influx into the cell. Accordingly, the compositions disclosed herein are believed to be useful for modulating calcium transport into and out of such intracellular stores and for the treatment of disorders that are characterized by aberrant calcium transport into and out of such intracellular stores. In particular, we believe that the SOC/CRAC calcium channel polypeptides disclosed herein play an important role in the influx of extracellular calcium by mediating the refilling of intracellular calcium stores following their depletion. Accordingly, we believe that the compositions for expressing functional SOC/CRAC calcium channel polypeptides in cells, as disclosed herein, are useful for treating patients having conditions that are characterized by reduced extracellular calcium influx into their SOC/CRAC-expressing cells. Additionally, the compositions of the invention are useful for delivering therapeutic and/or imaging agents to cells which preferentially express SOC/CRAC calcium channel polypeptides and, in particular, for delivering such agents to hematopoietic cells, liver, heart, spleen, and kidney to modulate proliferation and growth of these cells. Moreover, in view of the importance of cellular calcium levels to cell viability, we believe that SOC-2/CRAC-1, SOC-3/CRAC-2, and SOC-4/CRAC-3 as disclosed herein, and/or other members of the SOC/CRAC family of calcium channel polypeptides, represent an ideal target for designing and/or identifying (e.g., from molecular libraries) small molecule inhibitors that block lymphocyte proliferation, as well as other binding agents that selectively bind to SOC/CRAC polypeptides to which drugs or toxins can be conjugated for delivery to SOC/CRAC polypeptide expressing cells.

The invention is based, in part, on the molecular cloning and sequence analysis of the novel SOC/CRAC calcium channel molecules disclosed herein (also referred to as a "SOC-2/CRAC-1 molecule," a "SOC-3/CRAC-2 molecule," and/or "SOC-4/CRAC-3 molecule") that are predominantly expressed in human hematopoietic cells, liver, spleen, heart, and

kidney (SOC-2/CRAC-1), kidney and colon (SOC-3/CRAC-2), and prostate (SOC-4/CRAC-3 molecule). As used herein, a "SOC/CRAC molecule" embraces a "SOC/CRAC calcium channel nucleic acid" (or "SOC/CRAC nucleic acid") and a "SOC/CRAC calcium channel polypeptide" (or "SOC/CRAC polypeptide"). Homologs and alleles also are embraced within the meaning of a SOC/CRAC calcium channel molecule.

According to one aspect of the invention, isolated SOC/CRAC nucleic acids which code for one or more member(s) of the SOC/CRAC family of calcium channel polypeptides or unique fragments thereof are provided. The isolated nucleic acids refer to one or more of the following:

(a) nucleic acid molecules which hybridize under stringent conditions to a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;

(b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;

(c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and

(d) complements of (a), (b) or (c).

The invention in another aspect provides an isolated nucleic acid molecule selected from the group consisting of (a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:29, and SEQ ID NO:31, (b) complements of (a), provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to any sequence selected from a sequence group consisting of (1) sequences having the SEQ. ID NOS. or GenBank accession numbers of Table I, (2) complements of (1), and (3) fragments of (1) and (2).

According to yet another aspect of the invention, isolated SOC/CRAC polypeptides are provided. The isolated SOC/CRAC polypeptide molecules are encoded by one or more SOC/CRAC nucleic acid molecules of the invention. Preferably, the SOC/CRAC polypeptide contains one or more polypeptides selected from the group consisting of the polypeptides having SEQ. ID Nos. 2, 4, 6, 8, 24, 26, 28, 30, and 32. In other embodiments, the isolated polypeptide may be a fragment or variant of the foregoing SOC/CRAC polypeptide molecules of sufficient length to represent a sequence unique within the human genome, and identifying

with a polypeptide that functions as a calcium channel, provided that the fragment excludes a sequence of contiguous amino acids identified in Table II, and/or excludes a sequence of contiguous amino acids encoded for by a nucleic acid sequence identified in Table I. In another embodiment, immunogenic fragments of the polypeptide molecules described above are provided.

According to another aspect of the invention, isolated SOC/CRAC binding agents (e.g., polypeptides) are provided which selectively bind to a SOC/CRAC molecule (e.g., a SOC/CRAC polypeptide encoded by the isolated nucleic acid molecules of the invention). Preferably, the isolated binding agents selectively bind to a polypeptide which comprises the sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32, or unique fragments thereof. In the preferred embodiments, the isolated binding polypeptides include antibodies and fragments of antibodies (e.g., Fab, F(ab)₂, Fd and antibody fragments which include a CDR3 region which binds selectively to a SOC/CRAC polypeptide). Preferably, the antibodies for human therapeutic applications are human antibodies.

According to another aspect of the invention, a pharmaceutical composition containing a pharmaceutically effective amount of an isolated SOC/CRAC nucleic acid, an isolated SOC/CRAC polypeptide, or an isolated SOC/CRAC binding polypeptide in a pharmaceutically acceptable carrier also is provided. The pharmaceutical compositions are useful in accordance with therapeutic methods disclosed herein.

According to yet another aspect of the invention, a method for isolating a SOC/CRAC molecule is provided. The method involves:

a) contacting a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample that is believed to contain one or more SOC/CRAC molecules, under conditions to form a complex of the SOC/CRAC nucleic acid or the SOC/CRAC binding polypeptide and the SOC/CRAC molecule;

b) detecting the presence of the complex;

c) isolating the SOC/CRAC molecule from the complex; and

d) determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity. As used herein "SOC/CRAC calcium channel activity" refers to the transport of Ca²⁺ into and out of intracellular stores that is mediated by a SOC/CRAC

polypeptide. In general, the SOC/CRAC calcium channel activity is initiated by a reduction or depletion of intracellular calcium stores.

In certain embodiments, the SOC/CRAC nucleic acid is a SOC-2/CRAC-1 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 27, or complements thereof); in certain other
5 embodiments, the SOC/CRAC nucleic acid is a SOC-3/CRAC-2 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 29, or complements thereof); in further embodiments, the SOC/CRAC nucleic acid is a SOC-4/CRAC-3 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 31, or complements thereof). In yet other embodiments, the SOC/CRAC polypeptide is a SOC-2/CRAC-1 binding polypeptide (e.g., an antibody that selectively binds to a SOC-
10 2/CRAC-1 polypeptide). In yet further embodiments, the SOC/CRAC polypeptide is a SOC-3/CRAC-2 binding polypeptide (e.g., an antibody that selectively binds to a SOC-3/CRAC-2 polypeptide). In some embodiments, the SOC/CRAC polypeptide is a SOC-4/CRAC-3 binding polypeptide (e.g., an antibody that selectively binds to a SOC-4/CRAC-3 polypeptide). In the preferred embodiments, the isolated binding polypeptides include
15 antibodies and fragments of antibodies (e.g., Fab, F(ab)₂, Fd and antibody fragments which include a CDR3 region which binds selectively to a SOC-2/CRAC-1, to a SOC-3/CRAC-2, and/or to a SOC-4/CRAC-3 polypeptide). Preferably the isolated binding polypeptides or other binding agents selectively bind to a single SOC/CRAC molecule, i.e., are capable of distinguishing between different members of the SOC/CRAC family. Accordingly, one or
20 more SOC/CRAC binding agents can be contained in a single composition (e.g., a pharmaceutical composition) to identify multiple SOC/CRAC molecules *in vivo* or *in vitro*.

According to yet another aspect of the invention, a method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity is provided. The method involves:

25 a) contacting a SOC/CRAC polypeptide with a candidate agent suspected of modulating SOC/CRAC calcium channel activity, under conditions sufficient to allow the candidate agent to interact selectively with (e.g. bind to) the SOC/CRAC polypeptide;

b) detecting a Ca²⁺ concentration of step (b) associated with the SOC/CRAC calcium channel activity of the SOC/CRAC polypeptide in the presence of the candidate agent; and

30 c) comparing the Ca²⁺ concentration of step (b) with a control Ca²⁺ concentration of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC calcium channel activity.

According to another aspect of the invention, a method for identifying agents useful in the modulation of a SOC/CRAC polypeptide kinase activity is provided. The method involves:

a) contacting a SOC/CRAC polypeptide with kinase activity with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity;

b) detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and

c) comparing the kinase activity of step (b) with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC kinase activity. In some embodiments the SOC/CRAC polypeptide comprises amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24), or a fragment thereof that retains the kinase activity.

According to yet another aspect of the invention, a method for determining the level of expression of a SOC/CRAC polypeptide in a subject is provided. The method involves:

a) measuring the expression of a SOC/CRAC polypeptide in a test sample, and

b) comparing the measured expression of the SOC/CRAC polypeptide in the test sample to the expression of a SOC/CRAC polypeptide in a control containing a known level of expression to determine the level of SOC/CRAC expression in the subject. Expression is defined as SOC/CRAC mRNA expression or SOC/CRAC polypeptide expression. Various methods can be used to measure expression. The preferred embodiments of the invention utilize PCR and Northern blotting for measuring mRNA expression, and monoclonal or polyclonal SOC/CRAC antisera as reagents for measuring SOC/CRAC polypeptide expression. In preferred embodiments, the SOC/CRAC molecule (nucleic acid and/or polypeptide) is SOC-2/CRAC-1. In other preferred embodiments, the SOC/CRAC molecule is SOC-3/CRAC-2. In yet further preferred embodiments, the SOC/CRAC molecule is SOC-4/CRAC-3. In certain embodiments, the test samples include biopsy samples and biological fluids such as blood. The method is useful, e.g., for assessing the presence or absence or stage of a proliferative disorder in a subject.

The invention also contemplates kits comprising a package including assays for SOC/CRAC epitopes, SOC/CRAC nucleic acids, and instructions, and optionally related materials such as controls, for example, a number, color chart, or an epitope of the expression product of the foregoing isolated nucleic acid molecules of the invention for comparing, for

example, the level of SOC/CRAC polypeptides or SOC/CRAC nucleic acid forms (wild-type or mutant) in a test sample to the level in a control sample having a known amount of a SOC/CRAC nucleic acid or SOC/CRAC polypeptide. This comparison can be used to assess in a subject a risk of developing a cancer or the progression of a cancer. The kits may also include assays for other known genes, and expression products thereof, associated with, for example, proliferative disorders (e.g., BRCA, p53, etc.). In a preferred embodiment, the kit comprises a package containing: (a) a binding agent that selectively binds to an isolated nucleic acid of the invention or an expression product thereof to obtain a measured test value, (b) a control containing a known amount of a SOC/CRAC nucleic acid or a SOC/CRAC polypeptide to obtain a measured control value, and (c) instructions for comparing the measured test value to the measured control value to determine the amount of SOC/CRAC nucleic acid or expression product thereof in a sample.

The invention provides isolated nucleic acid molecules, unique fragments thereof, expression vectors containing the foregoing, and host cells containing the foregoing. The invention also provides isolated binding polypeptides and binding agents which bind such polypeptides, including antibodies, and pharmaceutical compositions containing any of the compositions of the invention. The foregoing can be used, *inter alia*, in the diagnosis or treatment of conditions characterized by the aberrant expression levels and/or the presence of mutant forms of a SOC/CRAC nucleic acid or polypeptide. The invention also provides methods for identifying agents that alter the function of the SOC/CRAC polypeptide.

These and other aspects of the invention, as well as various advantages and utilities, will be more apparent with reference to the detailed description of the preferred embodiments.

Brief Description of the Sequences

SEQ ID NO:1 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:2 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:1).

SEQ ID NO:3 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:4 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:3).

SEQ ID NO:5 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:6 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:5).

SEQ ID NO:7 is a partial nucleotide sequence of the mouse homologue (mSOC-2/CRAC-1) of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:8 is the predicted amino acid sequence of the translation product of the mSOC-2/CRAC-1 cDNA (SEQ ID NO:7).

5 SEQ ID NO:9 is the nucleotide sequence of the mouse MLSN-1 (SOC-1) cDNA.

SEQ ID NO:10 is the predicted amino acid sequence of the translation product of the mouse MLSN-1 (SOC-1) cDNA (SEQ ID NO:9).

SEQ ID NO:11 is the nucleotide sequence of a human calcium channel cDNA with GenBank Acc. no.: AB001535.

10 SEQ ID NO:12 is the predicted amino acid sequence of the translation product of the human calcium channel cDNA with GenBank Acc. no.: AB001535 (SEQ ID NO:11).

SEQ ID NO:13 is the amino acid sequence of a *C. Elegans* polypeptide at the c05c12.3 locus.

15 SEQ ID NO:14 is the amino acid sequence of a *C. Elegans* polypeptide at the F54D1 locus.

SEQ ID NO:15 is the amino acid sequence of a *C. Elegans* polypeptide at the t01H8 locus.

SEQ ID NO:16 is the nucleotide sequence of a mouse kidney cDNA with GenBank Acc. no.: AI226731.

20 SEQ ID NO:17 is the predicted amino acid sequence of the translation product of the mouse kidney cDNA with GenBank Acc. no.: AI226731 (SEQ ID NO:16).

SEQ ID NO:18 is the nucleotide sequence of a human brain cDNA with GenBank Acc. no.: H18835.

25 SEQ ID NO:19 is the predicted amino acid sequence of the translation product of the human brain cDNA with GenBank Acc. no.: H18835 (SEQ ID NO:18).

SEQ ID NO:20 is the nucleotide sequence of the human EST with GenBank Acc. no.: AA419592.

SEQ ID NO:21 is the nucleotide sequence of the human EST with GenBank Acc. no.: AA419407.

30 SEQ ID NO:22 is the nucleotide sequence of the mouse EST with GenBank Acc. no.: AI098310.

SEQ ID NO:23 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA that contains the SOC-2/CRAC-1 sequences of SEQ ID NO:1, SEQ ID NO:3, and SEQ ID NO:5.

SEQ ID NO:24 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:23).

SEQ ID NO:25 is a partial nucleotide sequence of the human SOC-3/CRAC-2 cDNA.

SEQ ID NO:26 is the predicted amino acid sequence of the translation product of human SOC-3/CRAC-2 cDNA (SEQ ID NO:25).

SEQ ID NO:27 is the full nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:28 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:27).

SEQ ID NO:29 is the full nucleotide sequence of the human SOC-3/CRAC-2 cDNA.

SEQ ID NO:30 is the predicted amino acid sequence of the translation product of human SOC-3/CRAC-2 cDNA (SEQ ID NO:29).

SEQ ID NO:31 is the full nucleotide sequence of the human SOC-4/CRAC-3 cDNA.

SEQ ID NO:32 is the predicted amino acid sequence of the translation product of human SOC-4/CRAC-3 cDNA (SEQ ID NO:31).

Brief Description of the Drawings

Figure 1 is a schematic depicting the intron/exon organization of the chicken SOC-2/CRAC-1 genomic sequence, as well as the putative transmembrane (TM) domains, and the targeting constructs utilized in the knockout experiments.

Detailed Description of the Invention

One aspect of the invention involves the partial cloning of cDNAs encoding members of a novel family of calcium channel polypeptides, referred to herein as "SOC/CRAC" (designated "SOC" or "CRAC" or "ICRAC", for Store Operated Channels or Calcium Release Activated Channels, or CECH). Although not intending to be bound to any particular mechanism or theory, we believe that a SOC/CRAC family member is a transmembrane calcium channel that modulates Ca^{2+} flux "into" and "out of" a cell; in certain instances it may be activated upon depletion of Ca^{2+} from intracellular calcium stores, allowing Ca^{2+} influx into the cell.

The first three isolated SOC/CRAC members disclosed herein, define a new family of calcium channels which is distinct from previously described calcium channels, such as voltage gated calcium channels, ryanodine receptor/inositol-1,4,5-triphosphate receptor

channels, and Transient Receptor Potential (TRP) channels. The SOC/CRAC family of calcium channels exhibits high selectivity (with a P_{Ca}/P_{Na} ratio near 1000), a unitary conductance below the detection level of the patch clamp method (the conductance estimated at approximately 0.2 picosiemens), and are subject to inhibition by high intracellular calcium levels. Although not intending to be bound to any particular mechanism or theory, we believe that SOC/CRAC calcium channels are responsible for the majority of, for example, calcium entry which occurs when intracellular calcium stores are depleted, and that SOC/CRAC currents are important for initiating various types of calcium-dependent processes. Thus, we believe that SOC/CRAC calcium channels play an important role in cellular calcium homeostasis by, e.g., modulating the supply of calcium to refill intracellular stores when depleted.

The isolated full-length sequence of a representative, first member of the SOC/CRAC family, human SOC/CRAC nucleic acid (cDNA), SOC-2/CRAC-1, is represented as the nucleic acid of SEQ ID NO:27. This nucleic acid sequence codes for the SOC-2/CRAC-1 polypeptide with the predicted amino acid sequence disclosed herein as SEQ ID NO:28. A homologous mouse cDNA sequence (>90% identity to the human at the nucleotide level) is represented as the nucleic acid of SEQ ID NO:7, and codes for a unique fragment of a mouse SOC-2/CRAC-1 polypeptide having the predicted, partial amino acid sequence represented as SEQ ID NO:8. Analysis of the SOC-2/CRAC-1 partial sequence by comparison to nucleic acid and protein databases show that SOC-2/CRAC-1 shares a limited homology to mouse MLSN-1 (SOC-1, SEQ ID NOs: 9 and 10). Limited homology is also shared between SOC-2/CRAC-1 and three *C. Elegans* polypeptides (SEQ ID NOs: 13, 14, and 15). We further believe that SOC-2/CRAC-1 plays a role in the regulation of cellular Ca^{2+} fluxing and, in particular, lymphocyte Ca^{2+} fluxing.

A second member of the human SOC/CRAC family of calcium channels, SOC-3/CRAC-2, is represented as the nucleic acid of SEQ ID NO:29, and codes for the human SOC-3/CRAC-2 polypeptide having the predicted amino acid sequence represented as SEQ ID NO:30 (this molecule may also be referred to as CECH2). SOC-3/CRAC-2 is predominantly expressed in human hematopoietic cells (including peripheral blood lymphocytes, liver, bone marrow, spleen, thymus, lymph nodes, heart, and kidney. Expression can also be detected (at lesser levels) in brain, skeletal muscle colon, small intestine, placenta, lung, and cells (cell lines) such as HL-60, HeLa, K562, MOLT-4, SW-480, A459, and G361.

A third member of the human SOC/CRAC family of calcium channels, SOC-4/CRAC-3, is represented as the nucleic acid of SEQ ID NO:31, and codes for the human SOC-4/CRAC-3 polypeptide having the predicted amino acid sequence represented as SEQ ID NO:32 (this molecule may also be referred to as CECH6). It specifically expressed in the prostate gland/cells.

As used herein, a SOC/CRAC calcium channel nucleic acid (also referred to herein as a "SOC/CRAC nucleic acid" refers to a nucleic acid molecule which: (1) hybridizes under stringent conditions to one or more of the nucleic acids having the sequences of SEQ. ID NOS. 7, 27, 29, and/or 31 (sequences of the mouse and human SOC-2/CRAC-1, human SOC-3/CRAC-2, and human SOC-4/CRAC-3 nucleic acids), and (2) codes for a SOC-2/CRAC-1, a SOC-3/CRAC-2 or a SOC-4/CRAC-3 calcium channel polypeptide, respectively, or unique fragments of said SOC-2/CRAC-1, SOC-3/CRAC-2, or SOC-4/CRAC-3 polypeptide.

As used herein, a SOC/CRAC calcium channel polypeptide (also referred to herein as a "SOC/CRAC polypeptide") refers to a polypeptide that is coded for by a SOC-2/CRAC-1, a SOC-3/CRAC-2, and/or a SOC-4/CRAC-3 nucleic acid. Preferably, the above-identified SOC/CRAC polypeptides mediate transport of calcium into and out of a cell.

SOC/CRAC polypeptides also are useful as immunogenic molecules for the generation of binding polypeptides (e.g., antibodies) which bind selectively to SOC/CRAC (e.g., SOC-2/CRAC-1, SOC-3/CRAC-2, and/or SOC-4/CRAC-3) polypeptides. Such antibodies can be used in diagnostic assays to identify and/or quantify the presence of a SOC/CRAC polypeptide in a sample, such as a biological fluid or biopsy sample. SOC/CRAC polypeptides further embrace functionally equivalent fragments, variants, and analogs of the preferred SOC/CRAC polypeptides, provided that the fragments, variants, and analogs also are useful in mediating calcium transport into and out of intracellular calcium stores.

As used herein, "SOC/CRAC calcium channel activity" refers to Ca^{2+} transport ("Ca²⁺ fluxing") across the plasma membrane that is mediated by a SOC/CRAC calcium channel polypeptide. The SOC/CRAC calcium channel polypeptide typically has one or more of the following properties: high selectivity, a unitary conductance below the detection level of the patch clamp method, and are subject to inhibition by high intracellular calcium levels. Such activity can be easily detected using standard methodology well known in the art. See, e.g., the Examples and Neher, E., "Ion channels for communication between and within cells",

Science, 1992; 256:498-502; and Hoth, M., and Penner, R., "Depletion of intracellular calcium stores activates a calcium current in mast cells", Nature, 1992; 355 (6358):353-6.

According to one aspect of the invention, isolated nucleic acid molecules which code for one or more member(s) of the SOC/CRAC family of calcium channel polypeptides are provided. The isolated nucleic acid molecules are selected from the following groups:

(a) nucleic acid molecules which hybridize under stringent conditions to one or more nucleic acid molecules selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;

(b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;

(c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and

(d) complements of (a), (b) or (c).

In certain embodiments, the isolated nucleic acid molecule comprises one or more of nucleotides 1-1212 of SEQ ID NO:1; nucleotides 1-739 of SEQ ID NO:3; nucleotides 1-1579 of SEQ ID NO:5; nucleotides 1-5117 of SEQ ID NO:23; the mouse homolog for SOC-2/CRAC-1 corresponding to SEQ ID NO:7; nucleotides 1-2180 of SEQ ID NO:25; nucleotides 382-5976 of SEQ ID NO:27; nucleotides 73-3714 of SEQ ID NO:29; and nucleotides 23-3434 of SEQ ID NO:31. In yet other embodiments, the isolated nucleic acid molecule comprises a molecule which encodes a polypeptide having one or more sequences selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32.

According to yet another aspect of the invention, an isolated nucleic acid molecule is provided which is selected from the group consisting of:

(a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, (of sufficient length to represent a sequence unique within the human genome); and (b) complements of (a), provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to a sequence in the prior art as represented by the sequence group consisting of: (1) sequences having the SEQ ID NOs or GenBank accession numbers of Table I, (2) complements of (1), and (3) fragments of (1) and (2).

In some embodiments, the sequence of contiguous nucleotides is selected from the group consisting of (1) at least two contiguous nucleotides nonidentical to the sequence group, (2) at least three contiguous nucleotides nonidentical to the sequence group, (3) at least four contiguous nucleotides nonidentical to the sequence group, (4) at least five contiguous nucleotides nonidentical to the sequence group, (5) at least six contiguous nucleotides nonidentical to the sequence group, (6) at least seven contiguous nucleotides nonidentical to the sequence group.

In other embodiments, the unique fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 40 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, 200 nucleotides, 1000 nucleotides and every integer length therebetween.

According to another aspect of the invention, expression vectors and host cells containing (e.g., transformed or transfected with) expression vectors comprising the nucleic acid molecules disclosed herein operably linked to a promoter are provided. In certain preferred embodiments, the host cells are eukaryotic cells.

The isolated nucleic acid molecules disclosed herein have various utilities, including their use as probes and primers to identify additional members of the SOC/CRAC family of calcium channels, as diagnostic reagents for identifying the presence of SOC/CRAC polypeptides in biological or other samples, and as agents for generating SOC/CRAC binding polypeptides (e.g., antibodies) that can be used as reagents in diagnostic and therapeutic assays to identify the presence, absence, and/or amounts of a SOC/CRAC nucleic acid or polypeptide in a biological or other sample.

As used herein with respect to nucleic acids, the term "isolated" means: (i) amplified *in vitro* by, for example, polymerase chain reaction (PCR); (ii) recombinantly produced by cloning; (iii) purified, as by cleavage and gel separation; or (iv) synthesized by, for example, chemical synthesis. An isolated nucleic acid is one which is readily manipulatable by recombinant DNA techniques well known in the art. Thus, a nucleotide sequence contained in a vector in which 5' and 3' restriction sites are known or for which polymerase chain reaction (PCR) primer sequences have been disclosed is considered isolated but a nucleic acid sequence existing in its native state in its natural host is not. An isolated nucleic acid may be substantially purified, but need not be. For example, a nucleic acid that is isolated within a cloning or expression vector is not pure in that it may comprise only a tiny percentage of the

material in the cell in which it resides. Such a nucleic acid is isolated, however, as the term is used herein because it is readily manipulatable by standard techniques known to those of ordinary skill in the art.

As used herein with respect to polypeptides (discussed below), the term "isolated" means separated from its native environment in sufficiently pure form so that it can be manipulated or used for any one of the purposes of the invention. Thus, isolated means sufficiently pure to be used (i) to raise and/or isolate antibodies, (ii) as a reagent in an assay, or (iii) for sequencing, etc.

Homologs and alleles of the SOC/CRAC nucleic acids of the invention can be identified by conventional techniques. Thus, an aspect of the invention is those nucleic acid sequences which code for SOC/CRAC polypeptides and which hybridize to a nucleic acid molecule selected from a group consisting of the nucleic acid of SEQ ID NO:1, the nucleic acid of SEQ ID NO:3, the nucleic acid of SEQ ID NO:5, the nucleic acid of SEQ ID NO:7, the nucleic acid of SEQ ID NO:23, the nucleic acid of SEQ ID NO:25, the nucleic acid of SEQ ID NO:27, the nucleic acid of SEQ ID NO:29, and the nucleic acid of SEQ ID NO:31, under stringent conditions. The term "stringent conditions" as used herein refers to parameters with which the art is familiar. Nucleic acid hybridization parameters may be found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. More specifically, stringent conditions, as used herein, refers, for example, to hybridization at 65°C in hybridization buffer (3.5 x SSC, 0.02% Ficoll, 0.02% polyvinyl pyrrolidone, 0.02% Bovine Serum Albumin, 2.5mM NaH₂PO₄(pH7), 0.5% SDS, 2mM EDTA). SSC is 0.15M sodium chloride/0.15M sodium citrate, pH7; SDS is sodium dodecyl sulphate; and EDTA is ethylenediaminetetracetic acid. After hybridization, the membrane upon which the DNA is transferred is washed at 2 x SSC at room temperature and then at 0.1 x SSC/0.1 x SDS at temperatures up to 68°C.

There are other conditions, reagents, and so forth which can be used, and would result in a similar degree of stringency. The skilled artisan will be familiar with such conditions, and thus they are not given here. It will be understood, however, that the skilled artisan will be able to manipulate the conditions in a manner to permit the clear identification of homologs and alleles of the SOC/CRAC nucleic acids of the invention. The skilled artisan also is familiar with the methodology for screening cells and libraries for expression of such

molecules which then are routinely isolated, followed by isolation of the pertinent nucleic acid molecule and sequencing.

In general homologs and alleles typically will share at least 40% nucleotide identity and/or at least 50% amino acid identity to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and/or SEQ ID NO:31, and SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, respectively. In some instances sequences will share at least 50% nucleotide identity and/or at least 65% amino acid identity and in still other instances sequences will share at least 60% nucleotide identity and/or at least 75% amino acid identity. The homology can be calculated using various, publicly available software tools developed by NCBI (Bethesda, Maryland) that can be obtained through the internet (<ftp://ncbi.nlm.nih.gov/pub/>). Exemplary tools include the BLAST system available at <http://www.ncbi.nlm.nih.gov>. Pairwise and ClustalW alignments (BLOSUM30 matrix setting) as well as Kyte-Doolittle hydropathic analysis can be obtained using the MacVector sequence analysis software (Oxford Molecular Group). Watson-Crick complements of the foregoing nucleic acids also are embraced by the invention.

In screening for SOC/CRAC related genes, such as homologs and alleles of SOC-2/CRAC-1 and/or SOC-3/CRAC-2, a Southern blot may be performed using the foregoing conditions, together with a radioactive probe. After washing the membrane to which the DNA is finally transferred, the membrane can be placed against X-ray film or a phosphorimager plate to detect the radioactive signal.

Given that the expression of the SOC/CRAC gene is prominent in certain human tissues (e.g., SOC-2/CRAC-1: lymphoid tissue/heart, SOC-3/CRAC-2: kidney/colon, SOC-4/CRAC-3: prostate), and given the teachings herein of partial human SOC/CRAC cDNA clones, full-length and other mammalian sequences corresponding to the human SOC/CRAC partial nucleic acid sequences can be isolated from, for example, a cDNA library prepared from one or more of the tissues in which SOC-2/CRAC-1 expression is prominent, SOC-3/CRAC-2 is prominent, and/or SOC-4/CRAC-3 expression is prominent, using standard colony hybridization techniques.

The invention also includes degenerate nucleic acids which include alternative codons to those present in the native materials. For example, serine residues are encoded by the codons TCA, AGT, TCC, TCG, TCT and AGC. Each of the six codons is equivalent for the purposes of encoding a serine residue. Thus, it will be apparent to one of ordinary skill in the

art that any of the serine-encoding nucleotide triplets may be employed to direct the protein synthesis apparatus, *in vitro* or *in vivo*, to incorporate a serine residue into an elongating SOC/CRAC polypeptide. Similarly, nucleotide sequence triplets which encode other amino acid residues include, but are not limited to: CCA, CCC, CCG and CCT (proline codons); CGA, CGC, CGG, CGT, AGA and AGG (arginine codons); ACA, ACC, ACG and ACT (threonine codons); AAC and AAT (asparagine codons); and ATA, ATC and ATT (isoleucine codons). Other amino acid residues may be encoded similarly by multiple nucleotide sequences. Thus, the invention embraces degenerate nucleic acids that differ from the biologically isolated nucleic acids in codon sequence due to the degeneracy of the genetic code.

The invention also provides isolated unique fragments of an isolated nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31. A unique fragment is one that is a 'signature' for the larger nucleic acid. For example, the unique fragment is long enough to assure that its precise sequence is not found in molecules within the human genome outside of the SOC/CRAC nucleic acids defined above (and human alleles). Those of ordinary skill in the art may apply no more than routine procedures to determine if a fragment is unique within the human genome.

Unique fragments, however, exclude fragments completely composed of the nucleotide sequences of any of GenBank accession numbers and SEQ ID NOs listed in Table I (SEQ ID NO:9, AB001535, AI226731, H18835, AA419592, AA261842, AA419407, AI098310, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853), or other previously published sequences as of the filing date of this application.

A fragment which is completely composed of the sequence described in the foregoing GenBank deposits and SEQ ID NO:9, is one which does not include any of the nucleotides unique to the sequences of the invention. Thus, a unique fragment must contain a nucleotide sequence other than the exact sequence of those in GenBank or fragments thereof. The difference may be an addition, deletion or substitution with respect to the GenBank sequence or it may be a sequence wholly separate from the GenBank sequence.

Unique fragments can be used as probes in Southern and Northern blot assays to identify such nucleic acids, or can be used in amplification assays such as those employing PCR. As known to those skilled in the art, large probes such as 200, 250, 300 or more nucleotides are preferred for certain uses such as Southern and Northern blots, while smaller fragments will be preferred for uses such as PCR. Unique fragments also can be used to produce fusion proteins for generating antibodies or determining binding of the polypeptide fragments, as demonstrated in the Examples, or for generating immunoassay components. Likewise, unique fragments can be employed to produce nonfused fragments of the SOC/CRAC polypeptides, useful, for example, in the preparation of antibodies, immunoassays or therapeutic applications. Unique fragments further can be used as antisense molecules to inhibit the expression of SOC/CRAC nucleic acids and polypeptides, respectively.

As will be recognized by those skilled in the art, the size of the unique fragment will depend upon its conservancy in the genetic code. Thus, some regions of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and complements thereof, will require longer segments to be unique while others will require only short segments, typically between 12 and 32 nucleotides long (e.g. 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32 bases) or more, up to the entire length of the disclosed sequence. As mentioned above, this disclosure intends to embrace each and every fragment of each sequence, beginning at the first nucleotide, the second nucleotide and so on, up to 8 nucleotides short of the end, and ending anywhere from nucleotide number 8, 9, 10 and so on for each sequence, up to the very last nucleotide, (provided the sequence is unique as described above). Virtually any segment of the region of SEQ ID NO:1 beginning at nucleotide 1 and ending at nucleotide 1212, or SEQ ID NO:3 beginning at nucleotide 1 and ending at nucleotide 739, or SEQ ID NO:5 beginning at nucleotide 1 and ending at nucleotide 1579, or SEQ ID NO:7 beginning at nucleotide 1 and ending at nucleotide 3532, or SEQ ID NO:23 beginning at nucleotide 1 and ending at nucleotide 5117, SEQ ID NO:25 beginning at nucleotide 1 and ending at nucleotide 2180, SEQ ID NO:27 beginning at nucleotide 1 and ending at nucleotide 7419, or SEQ ID NO:29 beginning at nucleotide 1 and ending at nucleotide 4061, or SEQ ID NO:31 beginning at nucleotide 1 and ending at nucleotide 4646, or complements thereof, that is 20 or more nucleotides in length will be unique. Those skilled in the art are well versed in methods for selecting such sequences, typically on the basis of the ability of the unique

fragment to selectively distinguish the sequence of interest from other sequences in the human genome of the fragment to those on known databases typically is all that is necessary, although *in vitro* confirmatory hybridization and sequencing analysis may be performed.

As mentioned above, the invention embraces antisense oligonucleotides that selectively bind to a nucleic acid molecule encoding a SOC/CRAC polypeptide, to decrease SOC/CRAC calcium channel activity. When using antisense preparations of the invention, slow intravenous administration is preferred.

As used herein, the term "antisense oligonucleotide" or "antisense" describes an oligonucleotide that is an oligoribonucleotide, oligodeoxyribonucleotide, modified oligoribonucleotide, or modified oligodeoxyribonucleotide which hybridizes under physiological conditions to DNA comprising a particular gene or to an mRNA transcript of that gene and, thereby, inhibits the transcription of that gene and/or the translation of that mRNA. The antisense molecules are designed so as to interfere with transcription or translation of a target gene upon hybridization with the target gene or transcript. Those skilled in the art will recognize that the exact length of the antisense oligonucleotide and its degree of complementarity with its target will depend upon the specific target selected, including the sequence of the target and the particular bases which comprise that sequence. It is preferred that the antisense oligonucleotide be constructed and arranged so as to bind selectively with the target under physiological conditions, i.e., to hybridize substantially more to the target sequence than to any other sequence in the target cell under physiological conditions. Based upon SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, or upon allelic or homologous genomic and/or cDNA sequences, one of skill in the art can easily choose and synthesize any of a number of appropriate antisense molecules for use in accordance with the present invention. In order to be sufficiently selective and potent for inhibition, such antisense oligonucleotides should comprise at least 10 and, more preferably, at least 15 consecutive bases which are complementary to the target, although in certain cases modified oligonucleotides as short as 7 bases in length have been used successfully as antisense oligonucleotides (Wagner et al., *Nat. Med.* 1(11):1116-1118, 1995). Most preferably, the antisense oligonucleotides comprise a complementary sequence of 20-30 bases. Although oligonucleotides may be chosen which are antisense to any region of the gene or mRNA transcripts, in preferred embodiments the antisense oligonucleotides correspond to N-terminal or 5' upstream sites such as translation initiation, transcription initiation or promoter sites. In

addition, 3'-untranslated regions may be targeted by antisense oligonucleotides. Targeting to mRNA splicing sites has also been used in the art but may be less preferred if alternative mRNA splicing occurs. In addition, the antisense is targeted, preferably, to sites in which mRNA secondary structure is not expected (see, e.g., Sainio et al., *Cell Mol. Neurobiol.* 14(5):439-457, 1994) and at which proteins are not expected to bind. Finally, although, SEQ ID No:1 discloses a cDNA sequence, one of ordinary skill in the art may easily derive the genomic DNA corresponding to this sequence. Thus, the present invention also provides for antisense oligonucleotides which are complementary to the genomic DNA corresponding to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31. Similarly, antisense to allelic or homologous SOC/CRAC cDNAs and genomic DNAs are enabled without undue experimentation.

In one set of embodiments, the antisense oligonucleotides of the invention may be composed of "natural" deoxyribonucleotides, ribonucleotides, or any combination thereof. That is, the 5' end of one native nucleotide and the 3' end of another native nucleotide may be covalently linked, as in natural systems, via a phosphodiester internucleoside linkage. These oligonucleotides may be prepared by art recognized methods which may be carried out manually or by an automated synthesizer. They also may be produced recombinantly by vectors.

In preferred embodiments, however, the antisense oligonucleotides of the invention also may include "modified" oligonucleotides. That is, the oligonucleotides may be modified in a number of ways which do not prevent them from hybridizing to their target but which enhance their stability or targeting or which otherwise enhance their therapeutic effectiveness.

The term "modified oligonucleotide" as used herein describes an oligonucleotide in which (1) at least two of its nucleotides are covalently linked via a synthetic internucleoside linkage (i.e., a linkage other than a phosphodiester linkage between the 5' end of one nucleotide and the 3' end of another nucleotide) and/or (2) a chemical group not normally associated with nucleic acids has been covalently attached to the oligonucleotide. Preferred synthetic internucleoside linkages are phosphorothioates, alkylphosphonates, phosphorodithioates, phosphate esters, alkylphosphonothioates, phosphoramidates, carbamates, carbonates, phosphate triesters, acetamides, carboxymethyl esters and peptides.

The term "modified oligonucleotide" also encompasses oligonucleotides with a covalently modified base and/or sugar. For example, modified oligonucleotides include

oligonucleotides having backbone sugars which are covalently attached to low molecular weight organic groups other than a hydroxyl group at the 3' position and other than a phosphate group at the 5' position. Thus modified oligonucleotides may include a 2'-O-alkylated ribose group. In addition, modified oligonucleotides may include sugars such as arabinose instead of ribose. The present invention, thus, contemplates pharmaceutical preparations containing modified antisense molecules that are complementary to and hybridizable with, under physiological conditions, nucleic acids encoding SOC/CRAC polypeptides, together with pharmaceutically acceptable carriers. Antisense oligonucleotides may be administered as part of a pharmaceutical composition. Such a pharmaceutical composition may include the antisense oligonucleotides in combination with any standard physiologically and/or pharmaceutically acceptable carriers which are known in the art. The compositions should be sterile and contain a therapeutically effective amount of the antisense oligonucleotides in a unit of weight or volume suitable for administration to a patient. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. The term "physiologically acceptable" refers to a non-toxic material that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art.

The invention also involves expression vectors coding for SOC/CRAC proteins and fragments and variants thereof and host cells containing those expression vectors. Virtually any cells, prokaryotic or eukaryotic, which can be transformed with heterologous DNA or RNA and which can be grown or maintained in culture, may be used in the practice of the invention. Examples include bacterial cells such as *E.coli* and eukaryotic cells such as mouse, hamster, pig, goat, primate, yeast, xenopous, etc. They may be of a wide variety of tissue types, including mast cells, fibroblasts, oocytes and lymphocytes, and they may be primary cells or cell lines. Specific examples include CHO cells and COS cells. Cell-free transcription systems also may be used in lieu of cells.

As used herein, a "vector" may be any of a number of nucleic acids into which a desired sequence may be inserted by restriction and ligation for transport between different genetic environments or for expression in a host cell. Vectors are typically composed of DNA although RNA vectors are also available. Vectors include, but are not limited to,

plasmids, phagemids and virus genomes. A cloning vector is one which is able to replicate in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence may be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence may occur many times as the plasmid increases in copy number within the host bacterium or just a single time per host before the host reproduces by mitosis. In the case of phage, replication may occur actively during a lytic phase or passively during a lysogenic phase. An expression vector is one into which a desired DNA sequence may be inserted by restriction and ligation such that it is operably joined to regulatory sequences and may be expressed as an RNA transcript. Vectors may further contain one or more marker sequences suitable for use in the identification of cells which have or have not been transformed or transfected with the vector. Markers include, for example, genes encoding proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art (e.g., β -galactosidase or alkaline phosphatase), and genes which visibly affect the phenotype of transformed or transfected cells, hosts, colonies or plaques (e.g., green fluorescent protein). Preferred vectors are those capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

As used herein, a coding sequence and regulatory sequences are said to be "operably" joined when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

The precise nature of the regulatory sequences needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 5' non-transcribed

and 5' non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. Especially, such 5' non-transcribed regulatory sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined gene.

5 Regulatory sequences may also include enhancer sequences or upstream activator sequences as desired. The vectors of the invention may optionally include 5' leader or signal sequences. The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

According to yet another aspect of the invention, isolated SOC/CRAC polypeptides

10 are provided. Preferably, the isolated SOC/CRAC polypeptides are encoded by the isolated SOC/CRAC nucleic acid molecules disclosed herein. More preferably, the isolated SOC/CRAC polypeptides of the invention are encoded by the nucleic acid molecules having SEQ ID Nos. 1, 3, 5, 7, 23, 25, 27, 29, and 31. In yet other embodiments, the isolated SOC/CRAC polypeptides of the invention have an amino acid sequence selected from the

15 group consisting of SEQ ID Nos. 2, 4, 6, 8, 24, 26, 28, 30 and 32. Preferably, the isolated SOC/CRAC polypeptides are of sufficient length to represent a sequence unique within the human genome. Thus, the preferred embodiments include a sequence of contiguous amino acids which is not identical to a prior art sequence as represented by the sequence group consisting of the contiguous amino acids identified in Table II (SEQ ID NO:10, SEQ ID

20 NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19 and GenBank Acc. Nos. AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572).

In certain embodiments, the isolated SOC/CRAC polypeptides are immunogenic and

25 can be used to generate binding polypeptides (e.g., antibodies) for use in diagnostic and therapeutic applications. Such binding polypeptides also are useful for detecting the presence, absence, and/or amounts of a SOC/CRAC nucleic acid or polypeptide in a sample such as a biological fluid or biopsy sample. Preferably, the SOC/CRAC polypeptides that are useful for generating binding polypeptides are unique polypeptides and, therefore, binding of the

30 antibody to a SOC/CRAC polypeptide in a sample is selective for the SOC/CRAC polypeptide.

Expression vectors containing all the necessary elements for expression are commercially available and known to those skilled in the art. See, e.g., Sambrook et al.,

Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989. Cells are genetically engineered by the introduction into the cells of heterologous DNA (RNA) encoding a SOC/CRAC polypeptide or fragment or variant thereof. The heterologous DNA (RNA) is placed under operable control of transcriptional elements to permit the expression of the heterologous DNA in the host cell.

Preferred systems for mRNA expression in mammalian cells are those such as pRc/CMV (available from Invitrogen, Carlsbad, CA) that contain a selectable marker such as a gene that confers G418 resistance (which facilitates the selection of stably transfected cell lines) and the human cytomegalovirus (CMV) enhancer-promoter sequences. Additionally, suitable for expression in primate or canine cell lines is the pCEP4 vector (Invitrogen, Carlsbad, CA), which contains an Epstein Barr virus (EBV) origin of replication, facilitating the maintenance of plasmid as a multicopy extrachromosomal element. Another expression vector is the pEF-BOS plasmid containing the promoter of polypeptide Elongation Factor 1 α , which stimulates efficiently transcription *in vitro*. The plasmid is described by Mishizuma and Nagata (*Nuc. Acids Res.* 18:5322, 1990), and its use in transfection experiments is disclosed by, for example, Demoulin (*Mol. Cell. Biol.* 16:4710-4716, 1996). Still another preferred expression vector is an adenovirus, described by Stratford-Perricaudet, which is defective for E1 and E3 proteins (*J. Clin. Invest.* 90:626-630, 1992). The use of the adenovirus as an Adeno.P1A recombinant is disclosed by Warnier et al., in intradermal injection in mice for immunization against P1A (*Int. J. Cancer*, 67:303-310, 1996).

The invention also embraces so-called expression kits, which allow the artisan to prepare a desired expression vector or vectors. Such expression kits include at least separate portions of each of the previously discussed coding sequences. Other components may be added, as desired, as long as the previously mentioned sequences, which are required, are included.

It will also be recognized that the invention embraces the use of the above described, SOC/CRAC cDNA sequence containing expression vectors, to transfect host cells and cell lines, by these prokaryotic (e.g., *E. coli*), or eukaryotic (e.g., CHO cells, COS cells, yeast expression systems and recombinant baculovirus expression in insect cells). Especially useful are mammalian cells such as mouse, hamster, pig, goat, primate, etc. They may be of a wide variety of tissue types, and include primary cells and cell lines. Specific examples include dendritic cells, U293 cells, peripheral blood leukocytes, bone marrow stem cells and embryonic stem cells. The invention also permits the construction of SOC/CRAC gene

“knock-outs” in cells and in animals, providing materials for studying certain aspects of SOC/CRAC calcium channel activity.

The invention also provides isolated polypeptides (including whole proteins and partial proteins), encoded by the foregoing SOC/CRAC nucleic acids, and include the polypeptides of SEQ ID NO:2, 4, 6, 8, 24, 26, 28, 30, 32, and unique fragments thereof. Such polypeptides are useful, for example, to regulate calcium transport-mediated cell growth, differentiation and proliferation, to generate antibodies, as components of immunoassays, etc. Polypeptides can be isolated from biological samples including tissue or cell homogenates, and can also be expressed recombinantly in a variety of prokaryotic and eukaryotic expression systems by constructing an expression vector appropriate to the expression system, introducing the expression vector into the expression system, and isolating the recombinantly expressed protein. Short polypeptides, including antigenic peptides (such as are presented by MHC molecules on the surface of a cell for immune recognition) also can be synthesized chemically using well-established methods of peptide synthesis.

A unique fragment of a SOC/CRAC polypeptide, in general, has the features and characteristics of unique fragments as discussed above in connection with nucleic acids. As will be recognized by those skilled in the art, the size of the unique fragment will depend upon factors such as whether the fragment constitutes a portion of a conserved protein domain. Thus, some regions of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, will require longer segments to be unique while others will require only short segments, typically between 5 and 12 amino acids (e.g. 5, 6, 7, 8, 9, 10, 11 and 12 amino acids long or more, including each integer up to the full length, >1,000 amino acids long). Virtually any segment of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, excluding the ones that share identity with it (the polypeptides identified in Table II - SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, and GenBank Acc. Nos. AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572) that is 9 or more amino acids in length will be unique.

Unique fragments of a polypeptide preferably are those fragments which retain a distinct functional capability of the polypeptide. Functional capabilities which can be retained in a unique fragment of a polypeptide include Ca^{2+} fluxing, high selectivity, a unitary

conductance below the detection level of the patch clamp method, and/or and are subject to inhibition by high intracellular calcium levels.

One important aspect of a unique fragment is its ability to act as a signature for identifying the polypeptide. Optionally, another aspect of a unique fragment is its ability to provide an immune response in an animal. Those skilled in the art are well versed in methods for selecting unique amino acid sequences, typically on the basis of the ability of the unique fragment to selectively distinguish the sequence of interest from non-family members. A comparison of the sequence of the fragment to those on known databases typically is all that is necessary.

The invention embraces variants of the SOC/CRAC polypeptides described above. As used herein, a "variant" of a SOC/CRAC polypeptide is a polypeptide which contains one or more modifications to the primary amino acid sequence of a SOC/CRAC polypeptide. Modifications which create a SOC/CRAC polypeptide variant are typically made to the nucleic acid which encodes the SOC/CRAC polypeptide, and can include deletions, point mutations, truncations, amino acid substitutions and addition of amino acids or non-amino acid moieties to: 1) reduce or eliminate a calcium channel activity of a SOC/CRAC polypeptide; 2) enhance a property of a SOC/CRAC polypeptide, such as protein stability in an expression system or the stability of protein-protein binding; 3) provide a novel activity or property to a SOC/CRAC polypeptide, such as addition of an antigenic epitope or addition of a detectable moiety; or 4) to provide equivalent or better binding to a SOC/CRAC polypeptide receptor or other molecule. Alternatively, modifications can be made directly to the polypeptide, such as by cleavage, addition of a linker molecule, addition of a detectable moiety, such as biotin, addition of a fatty acid, and the like. Modifications also embrace fusion proteins comprising all or part of the SOC/CRAC amino acid sequence. One of skill in the art will be familiar with methods for predicting the effect on protein conformation of a change in protein sequence, and can thus "design" a variant SOC/CRAC polypeptide according to known methods. One example of such a method is described by Dahiyat and Mayo in *Science* 278:82-87, 1997, whereby proteins can be designed *de novo*. The method can be applied to a known protein to vary only a portion of the polypeptide sequence. By applying the computational methods of Dahiyat and Mayo, specific variants of a SOC/CRAC calcium channel polypeptide can be proposed and tested to determine whether the variant retains a desired conformation.

5 Variants can include SOC/CRAC polypeptides which are modified specifically to alter a feature of the polypeptide unrelated to its physiological activity. For example, cysteine residues can be substituted or deleted to prevent unwanted disulfide linkages. Similarly, certain amino acids can be changed to enhance expression of a SOC/CRAC polypeptide by eliminating proteolysis by proteases in an expression system (e.g., dibasic amino acid residues in yeast expression systems in which KEX2 protease activity is present).

10 Mutations of a nucleic acid which encodes a SOC/CRAC polypeptide preferably preserve the amino acid reading frame of the coding sequence and, preferably, do not create regions in the nucleic acid which are likely to hybridize to form secondary structures, such as hairpins or loops, which can be deleterious to expression of the variant polypeptide.

15 Mutations can be made by selecting an amino acid substitution, or by random mutagenesis of a selected site in a nucleic acid which encodes the polypeptide. Variant polypeptides are then expressed and tested for one or more activities to determine which mutation provides a variant polypeptide with the desired properties. Further mutations can be made to variants (or to non-variant SOC/CRAC polypeptides) which are silent as to the amino acid sequence of the polypeptide, but which provide preferred codons for translation in a particular host. The preferred codons for translation of a nucleic acid in, e.g., E. coli, are well known to those of ordinary skill in the art. Still other mutations can be made to the noncoding sequences of a SOC/CRAC gene or cDNA clone to enhance expression of the polypeptide.

20 The skilled artisan will realize that conservative amino acid substitutions may be made in SOC/CRAC polypeptides to provide functionally equivalent variants of the foregoing polypeptides, i.e., the variants retain the functional capabilities of the SOC/CRAC polypeptides. As used herein, a "conservative amino acid substitution" refers to an amino acid substitution which does not alter the relative charge or size characteristics of the protein in which the amino acid substitution is made. Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Exemplary functionally equivalent variants of the SOC/CRAC polypeptides include conservative amino acid substitutions of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32. Conservative substitutions of amino acids

25

30

include substitutions made amongst amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D.

Thus functionally equivalent variants of SOC/CRAC polypeptides, i.e., variants of SOC/CRAC polypeptides which retain the function of the natural SOC/CRAC polypeptides, are contemplated by the invention. Conservative amino-acid substitutions in the amino acid sequence of SOC/CRAC polypeptides to produce functionally equivalent variants of SOC/CRAC polypeptides typically are made by alteration of a nucleic acid encoding SOC/CRAC polypeptides (e.g., SEQ ID NOs:1, 3, 5, 7, 23, 25, 27, 29, 31). Such substitutions can be made by a variety of methods known to one of ordinary skill in the art. For example, amino acid substitutions may be made by PCR-directed mutation, site-directed mutagenesis according to the method of Kunkel (Kunkel, *Proc. Nat. Acad. Sci. U.S.A.* 82: 488-492, 1985), or by chemical synthesis of a gene encoding a SOC/CRAC polypeptide. The activity of functionally equivalent fragments of SOC/CRAC polypeptides can be tested by cloning the gene encoding the altered SOC/CRAC polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the altered SOC/CRAC polypeptide, and testing for a functional capability of the SOC/CRAC polypeptides as disclosed herein (e.g., SOC/CRAC calcium channel activity).

The invention as described herein has a number of uses, some of which are described elsewhere herein. First, the invention permits isolation of SOC/CRAC polypeptides, including the isolation of the complete SOC/CRAC polypeptide. A variety of methodologies well-known to the skilled practitioner can be utilized to obtain isolated SOC/CRAC molecules. The polypeptide may be purified from cells which naturally produce the polypeptide by chromatographic means or immunological recognition. Alternatively, an expression vector may be introduced into cells to cause production of the polypeptide. In another method, mRNA transcripts may be microinjected or otherwise introduced into cells to cause production of the encoded polypeptide. Translation of SOC/CRAC mRNA in cell-free extracts such as the reticulocyte lysate system also may be used to produce SOC/CRAC polypeptides. Those skilled in the art also can readily follow known methods for isolating SOC/CRAC polypeptides. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography and immune-affinity chromatography.

The invention also provides, in certain embodiments, "dominant negative" polypeptides derived from SOC/CRAC polypeptides. A dominant negative polypeptide is an

inactive variant of a protein, which, by interacting with the cellular machinery, displaces an active protein from its interaction with the cellular machinery or competes with the active protein, thereby reducing the effect of the active protein. For example, a dominant negative receptor which binds a ligand but does not transmit a signal in response to binding of the ligand can reduce the biological effect of expression of the ligand. Likewise, a dominant negative inactive SOC/CRAC calcium channel which interacts normally with the cell membrane but which does not mediate calcium transport can reduce calcium transport in a cell. Similarly, a dominant negative transcription factor which binds to a promoter site in the control region of a gene but does not increase gene transcription can reduce the effect of a normal transcription factor by occupying promoter binding sites without increasing transcription.

The end result of the expression of a dominant negative polypeptide in a cell is a reduction in function of active proteins. One of ordinary skill in the art can assess the potential for a dominant negative variant of a protein, and using standard mutagenesis techniques to create one or more dominant negative variant polypeptides. See, e.g., U.S. Patent No. 5,580,723 and Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press. The skilled artisan then can test the population of mutagenized polypeptides for diminution in a selected and/or for retention of such an activity. Other similar methods for creating and testing dominant negative variants of a protein will be apparent to one of ordinary skill in the art.

According to another aspect, the invention provides a method for isolating a SOC/CRAC molecule having SOC/CRAC calcium channel activity. The method involves contacting a binding molecule that is a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing one or more SOC/CRAC molecules under conditions that allow such binding (see earlier discussion) to form a complex, detecting the presence of the complex, isolating the SOC/CRAC molecule from the complex, and determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity. Thus, the invention is useful for identifying and isolating full length complementary (cDNA) or genomic nucleic acids encoding SOC/CRAC polypeptides having SOC/CRAC calcium channel activity. Identification and isolation of such nucleic acids and polypeptides may be accomplished by hybridizing/binding, under appropriate conditions well known in the art, libraries and/or restriction enzyme-digested human nucleic acids, with a labeled SOC/CRAC molecular probe. As used herein, a "label" includes molecules that are incorporated into, for

example, a SOC/CRAC molecule (nucleic acid or peptide), that can be directly or indirectly detected. A wide variety of detectable labels are well known in the art that can be used, and include labels that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc), or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseshoe peroxidase, etc.). The label may be bound to a SOC/CRAC binding partner, or incorporated into the structure of the binding partner.

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly detected through optical or electron density, radioactive emissions, nonradioactive energy transfers, etc. or indirectly detected with antibody conjugates, streptavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art. Once a library clone or hybridizing fragment is identified in the hybridization/binding reaction, it can be further isolated by employing standard isolation/cloning techniques known to those of skill in the art. See, generally, Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press. In addition, nucleic acid amplification techniques well known in the art, may also be used to locate splice variants of calcium channel (or calcium channel subunits) with SOC/CRAC calcium channel activity. Size and sequence determinations of the amplification products can reveal splice variants.

The foregoing isolated nucleic acids and polypeptides may then be compared to the nucleic acids and polypeptides of the present invention in order to identify homogeneity or divergence of the sequences, and be further characterized functionally to determine whether they belong to a family of molecules with SOC/CRAC calcium channel activity (for methodology see under the Examples section).

The isolation of the SOC/CRAC cDNA and/or partial sequences thereof also makes it possible for the artisan to diagnose a disorder characterized by an aberrant expression of SOC/CRAC. These methods involve determining expression of the SOC/CRAC gene, and/or SOC/CRAC polypeptides derived therefrom. In the former situation, such determinations can be carried out via any standard nucleic acid determination assay, including the polymerase chain reaction, or assaying with labeled hybridization probes as exemplified below. In the latter situation, such determination can be carried out via any standard immunological assay using, for example, antibodies which bind to the SOC/CRAC protein.

The invention also embraces isolated peptide binding agents which, for example, can be antibodies or fragments of antibodies ("binding polypeptides"), having the ability to selectively bind to SOC/CRAC polypeptides. Antibodies include polyclonal and monoclonal antibodies, prepared according to conventional methodology. In certain embodiments, the invention excludes binding agents (e.g., antibodies) that bind to the polypeptides encoded by the nucleic acids of SEQ ID NOs: 10, 12, 13, 14, 15, 17, and 19.

Significantly, as is well-known in the art, only a small portion of an antibody molecule, the paratope, is involved in the binding of the antibody to its epitope (see, in general, Clark, W.R. (1986) The Experimental Foundations of Modern Immunology Wiley & Sons, Inc., New York; Roitt, I. (1991) Essential Immunology, 7th Ed., Blackwell Scientific Publications, Oxford). The pFc' and Fc regions, for example, are effectors of the complement cascade but are not involved in antigen binding. An antibody from which the pFc' region has been enzymatically cleaved, or which has been produced without the pFc' region, designated an F(ab')₂ fragment, retains both of the antigen binding sites of an intact antibody. Similarly, an antibody from which the Fc region has been enzymatically cleaved, or which has been produced without the Fc region, designated an Fab fragment, retains one of the antigen binding sites of an intact antibody molecule. Proceeding further, Fab fragments consist of a covalently bound antibody light chain and a portion of the antibody heavy chain denoted Fd. The Fd fragments are the major determinant of antibody specificity (a single Fd fragment may be associated with up to ten different light chains without altering antibody specificity) and Fd fragments retain epitope-binding ability in isolation.

Within the antigen-binding portion of an antibody, as is well-known in the art, there are complementarity determining regions (CDRs), which directly interact with the epitope of the antigen, and framework regions (FRs), which maintain the tertiary structure of the paratope (see, in general, Clark, 1986; Roitt, 1991). In both the heavy chain Fd fragment and the light chain of IgG immunoglobulins, there are four framework regions (FR1 through FR4) separated respectively by three complementarity determining regions (CDR1 through CDR3). The CDRs, and in particular the CDR3 regions, and more particularly the heavy chain CDR3, are largely responsible for antibody specificity.

It is now well-established in the art that the non-CDR regions of a mammalian antibody may be replaced with similar regions of conspecific or heterospecific antibodies while retaining the epitopic specificity of the original antibody. This is most clearly manifested in the development and use of "humanized" antibodies in which non-human CDRs

are covalently joined to human FR and/or Fc/pFc' regions to produce a functional antibody. Thus, for example, PCT International Publication Number WO 92/04381 teaches the production and use of humanized murine RSV antibodies in which at least a portion of the murine FR regions have been replaced by FR regions of human origin. Such antibodies, including fragments of intact antibodies with antigen-binding ability, are often referred to as "chimeric" antibodies.

Thus, as will be apparent to one of ordinary skill in the art, the present invention also provides for F(ab')₂, Fab, Fv and Fd fragments; chimeric antibodies in which the Fc and/or FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric F(ab')₂ fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric Fab fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; and chimeric Fd fragment antibodies in which the FR and/or CDR1 and/or CDR2 regions have been replaced by homologous human or non-human sequences. The present invention also includes so-called single chain antibodies.

Thus, the invention involves binding polypeptides of numerous size and type that bind selectively to SOC/CRAC polypeptides, and complexes containing SOC/CRAC polypeptides. These binding polypeptides also may be derived also from sources other than antibody technology. For example, such polypeptide binding agents can be provided by degenerate peptide libraries which can be readily prepared in solution, in immobilized form, as bacterial flagella peptide display libraries or as phage display libraries. Combinatorial libraries also can be synthesized of peptides containing one or more amino acids. Libraries further can be synthesized of peptides and non-peptide synthetic moieties.

Phage display can be particularly effective in identifying binding peptides useful according to the invention. Briefly, one prepares a phage library (using e.g. m13, fd, or lambda phage), displaying inserts from 4 to about 80 amino acid residues using conventional procedures. The inserts may represent, for example, a completely degenerate or biased array. One then can select phage-bearing inserts which bind to the SOC/CRAC polypeptide or a complex containing a SOC/CRAC polypeptide. This process can be repeated through several cycles of reselection of phage that bind to the SOC/CRAC polypeptide or complex. Repeated rounds lead to enrichment of phage bearing particular sequences. DNA sequence analysis can be conducted to identify the sequences of the expressed polypeptides. The minimal linear

portion of the sequence that binds to the SOC/CRAC polypeptide or complex can be determined. One can repeat the procedure using a biased library containing inserts containing part or all of the minimal linear portion plus one or more additional degenerate residues upstream or downstream thereof. Yeast two-hybrid screening methods also may be used to
5 identify polypeptides that bind to the SOC/CRAC polypeptides. Thus, the SOC/CRAC polypeptides of the invention, or a fragment thereof, or complexes of SOC/CRAC can be used to screen peptide libraries, including phage display libraries, to identify and select peptide binding polypeptides that selectively bind to the SOC/CRAC polypeptides of the invention. Such molecules can be used, as described, for screening assays, for purification protocols, for
10 interfering directly with the functioning of SOC/CRAC and for other purposes that will be apparent to those of ordinary skill in the art.

A SOC/CRAC polypeptide, or a fragment thereof, also can be used to isolate naturally occurring, polypeptide binding partners which may associate with the SOC/CRAC polypeptide in the membrane of a cell. Isolation of binding partners may be performed
15 according to well-known methods. For example, isolated SOC/CRAC polypeptides can be attached to a substrate, and then a solution suspected of containing an SOC/CRAC binding partner may be applied to the substrate. If the binding partner for SOC/CRAC polypeptides is present in the solution, then it will bind to the substrate-bound SOC/CRAC polypeptide. The binding partner then may be isolated. Other proteins which are binding partners for
20 SOC/CRAC, may be isolated by similar methods without undue experimentation.

The invention also provides novel kits which could be used to measure the levels of the nucleic acids of the invention, expression products of the invention or anti-SOC/CRAC antibodies. In the case of nucleic acid detection, pairs of primers for amplifying SOC/CRAC nucleic acids can be included. The preferred kits would include controls such as known
25 amounts of nucleic acid probes, SOC/CRAC epitopes (such as SOC/CRAC expression products) or anti-SOC/CRAC antibodies, as well as instructions or other printed material. In certain embodiments the printed material can characterize risk of developing a disorder that is characterized by aberrant SOC/CRAC polypeptide expression based upon the outcome of the assay. The reagents may be packaged in containers and/or coated on wells in predetermined
30 amounts, and the kits may include standard materials such as labeled immunological reagents (such as labeled anti-IgG antibodies) and the like. One kit is a packaged polystyrene microtiter plate coated with a SOC/CRAC polypeptide and a container containing labeled anti-human IgG antibodies. A well of the plate is contacted with, for example, serum, washed

and then contacted with the anti-IgG antibody. The label is then detected. A kit embodying features of the present invention is comprised of the following major elements: packaging an agent of the invention, a control agent, and instructions. Packaging is a box-like structure for holding a vial (or number of vials) containing an agent of the invention, a vial (or number of vials) containing a control agent, and instructions. Individuals skilled in the art can readily modify packaging to suit individual needs.

Another aspect of the invention is a method for determining the level of SOC/CRAC expression in a subject. As used herein, a subject is a human, non-human primate, cow, horse, pig, sheep, goat, dog, cat or rodent. In all embodiments, human subjects are preferred. Expression is defined either as SOC/CRAC mRNA expression or SOC/CRAC polypeptide expression. Various methods can be used to measure expression. Preferred embodiments of the invention include PCR and Northern blotting for measuring mRNA expression, and monoclonal or polyclonal SOC/CRAC antisera as reagents to measure SOC/CRAC polypeptide expression. In certain embodiments, test samples such as biopsy samples, and biological fluids such as blood, are used as test samples. SOC/CRAC expression in a test sample of a subject is compared to SOC/CRAC expression in control sample to, e.g., assess the presence or absence or stage of a proliferative disorder (e.g., a lymphocyte proliferative disorder) in a subject.

SOC/CRAC polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts. Recombinantly produced SOC/CRAC polypeptides include chimeric proteins comprising a fusion of a SOC/CRAC protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the SOC/CRAC polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein. A polypeptide fused to a SOC/CRAC polypeptide or fragment may also provide means of readily detecting the fusion protein, e.g., by immunological recognition or by fluorescent labeling.

The invention is also useful in the generation of transgenic non-human animals. As used herein, "transgenic non-human animals" includes non-human animals having one or more exogenous nucleic acid molecules incorporated in germ line cells and/or somatic cells. Thus the transgenic animal include "knockout" animals having a homozygous or heterozygous gene disruption by homologous recombination, animals having episomal or chromosomally incorporated expression vectors, etc. Knockout animals can be prepared by

homologous recombination using embryonic stem cells as is well known in the art. The recombination may be facilitated using, for example, the cre/lox system or other recombinase systems known to one of ordinary skill in the art. In certain embodiments, the recombinase system itself is expressed conditionally, for example, in certain tissues or cell types, at certain embryonic or post-embryonic developmental stages, inducibly by the addition of a compound which increases or decreases expression, and the like. In general, the conditional expression vectors used in such systems use a variety of promoters which confer the desired gene expression pattern (e.g., temporal or spatial). Conditional promoters also can be operably linked to SOC/CRAC nucleic acid molecules to increase expression of SOC/CRAC in a regulated or conditional manner. *Trans*-acting negative regulators of SOC/CRAC calcium channel activity or expression also can be operably linked to a conditional promoter as described above. Such *trans*-acting regulators include antisense SOC/CRAC nucleic acids molecules, nucleic acid molecules which encode dominant negative SOC/CRAC molecules, ribozyme molecules specific for SOC/CRAC nucleic acids, and the like. The transgenic non-human animals are useful in experiments directed toward testing biochemical or physiological effects of diagnostics or therapeutics for conditions characterized by increased or decreased SOC/CRAC expression. Other uses will be apparent to one of ordinary skill in the art.

The invention further provides efficient methods of identifying agents or lead compounds for agents active at the level of a SOC/CRAC polypeptide (e.g., a SOC/CRAC polypeptide) or SOC/CRAC fragment dependent cellular function. In particular, such functions include interaction with other polypeptides or fragments thereof, and selective binding to certain molecules (e.g., agonists and antagonists). Generally, the screening methods involve assaying for compounds which interfere with SOC/CRAC calcium channel activity, although compounds which enhance SOC/CRAC calcium channel activity also can be assayed using the screening methods. Such methods are adaptable to automated, high throughput screening of compounds. The target therapeutic indications for pharmacological agents detected by the screening methods are limited only in that the target cellular function be subject to modulation by alteration of the formation of a complex comprising a SOC/CRAC polypeptide or fragment thereof and one or more SOC/CRAC binding targets. Target indications include cellular processes modulated by SOC/CRAC such as Ca^{2+} fluxing, and affected by SOC/CRAC ability to form complexes with other molecules and polypeptides as, for example, may be present in the cell membrane.

A wide variety of assays for pharmacological agents are provided, including, expression assays, labeled *in vitro* protein-protein binding assays, electrophoretic mobility shift assays, immunoassays, cell-based assays such as calcium transport assays, etc. For example, two-hybrid screens are used to rapidly examine the effect of transfected nucleic acids on the intracellular binding of SOC/CRAC or SOC/CRAC fragments to specific intracellular targets (e.g. a tyrosine kinase). The transfected nucleic acids can encode, for example, combinatorial peptide libraries or cDNA libraries. Convenient reagents for such assays, e.g., GAL4 fusion proteins, are known in the art. An exemplary cell-based assay involves transfecting a cell with a nucleic acid encoding a SOC/CRAC polypeptide fused to a GAL4 DNA binding domain and a nucleic acid encoding a reporter gene operably linked to a gene expression regulatory region, such as one or more GAL4 binding sites. Activation of reporter gene transcription occurs when the SOC/CRAC and reporter fusion polypeptides bind such as to enable transcription of the reporter gene. Agents which modulate a SOC/CRAC polypeptide mediated cell function are then detected through a change in the expression of reporter gene. Methods for determining changes in the expression of a reporter gene are known in the art.

In an expression system, for example, a SOC/CRAC polypeptide is attached to a membrane, the membrane preferably separating two fluid environments and being otherwise not permeable to Ca^{2+} . Such separation is preferred so that a change in Ca^{2+} concentration on either side of the membrane is mediated only through the attached SOC/CRAC polypeptide. Preferably, a SOC/CRAC polypeptide is expressed in an intact cell and is present on the cell-membrane (as in physiologic conditions). The cell expressing the SOC/CRAC polypeptide is preferably a eukaryotic cell, and the SOC/CRAC polypeptide is preferably recombinantly expressed, although cells naturally expressing a SOC/CRAC polypeptide may also be used. Synthetic membranes, however, containing SOC/CRAC polypeptides may also be used. See, e.g., K. Kiselyov, et al., Functional interaction between InsP3 receptors and store-operated Htrp3 channels, Nature 396, 478-82 (1998).

The cell expressing the SOC/CRAC polypeptide is incubated under conditions which, in the absence of the candidate agent, permit calcium flux into the cell and allow detection of a reference calcium concentration. For example, depletion of intracellular calcium stores with thapsigargin or other agents (Putney, J.W. Jr., in Capacitative Calcium Entry, R.G. Landes Co. and Chapman & Hall, 1997) would produce a given level of SOC/CRAC channel activation and a given reference calcium concentration. Detection of a decrease in the

foregoing activities (i.e., a decrease in the intracellular calcium concentration) relative to the reference calcium concentration indicates that the candidate agent is a lead compound for an agent to inhibit SOC/CRAC calcium channel activity. Preferred SOC/CRAC polypeptides include the polypeptides of claim 15.

5 SOC/CRAC fragments used in the methods, when not produced by a transfected nucleic acid are added to an assay mixture as an isolated polypeptide. SOC/CRAC polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts or chemically synthesized. Recombinantly produced SOC/CRAC polypeptides include chimeric proteins comprising a fusion of a SOC/CRAC
10 protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the SOC/CRAC polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein or Flag epitope.

The assay mixture is comprised of a SOC/CRAC polypeptide binding target
15 (candidate agent) capable of interacting with a SOC/CRAC polypeptide. While natural SOC/CRAC binding targets may be used, it is frequently preferred to use portions (e.g., peptides or nucleic acid fragments) or analogs (i.e., agents which mimic the SOC/CRAC binding properties of the natural binding target for purposes of the assay) of the SOC/CRAC binding target so long as the portion or analog provides binding affinity and avidity to the
20 SOC/CRAC polypeptide (or fragment thereof) measurable in the assay.

The assay mixture also comprises a candidate agent (binding target, e.g., agonist/antagonist). Typically, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a different response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration of agent or
25 at a concentration of agent below the limits of assay detection. Candidate agents encompass numerous chemical classes, although typically they are organic compounds. Preferably, the candidate agents are small organic compounds, i.e., those having a molecular weight of more than 50 yet less than about 2500, preferably less than about 1000 and, more preferably, less than about 500. Candidate agents comprise functional chemical groups necessary for
30 structural interactions with polypeptides and/or nucleic acids, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups and more preferably at least three of the functional chemical groups. The candidate agents can comprise cyclic carbon or heterocyclic structure and/or aromatic or

polyaromatic structures substituted with one or more of the above-identified functional groups. Candidate agents also can be biomolecules such as peptides, saccharides, fatty acids, sterols, isoprenoids, purines, pyrimidines, derivatives or structural analogs of the above, or combinations thereof and the like. Where the agent is a nucleic acid, the agent typically is a DNA or RNA molecule, although modified nucleic acids as defined herein are also contemplated.

Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides, synthetic organic combinatorial libraries, phage display libraries of random peptides, and the like. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural and synthetically produced libraries and compounds can be readily modified through conventional chemical, physical, and biochemical means. Further, known agents may be subjected to directed or random chemical modifications such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs of the agents. Non-SOC/CRAC calcium channel agonists and antagonists, for example, include agents such as dihydropyridines (DHPs), phenylalkylamines, omega conotoxin (omega-CgTx) and pyrazonoylguanidines.

A variety of other reagents also can be included in the mixture. These include reagents such as salts, buffers, neutral proteins (e.g., albumin), detergents, etc. which may be used to facilitate optimal protein-protein, protein-nucleic acid, and/or protein/membrane component binding association. Such a reagent may also reduce non-specific or background interactions of the reaction components. Other reagents that improve the efficiency of the assay such as protease, inhibitors, nuclease inhibitors, antimicrobial agents, and the like may also be used.

The mixture of the foregoing assay materials is incubated under conditions whereby, but for the presence of the candidate agent, the SOC/CRAC polypeptide specifically binds the cellular binding target, a portion thereof or analog thereof. The order of addition of components, incubation temperature, time of incubation, and other perimeters of the assay may be readily determined. Such experimentation merely involves optimization of the assay parameters, not the fundamental composition of the assay. Incubation temperatures typically

are between 4°C and 40°C. Incubation times preferably are minimized to facilitate rapid, high throughput screening, and typically are between 0.1 and 10 hours.

After incubation, the presence or absence of specific binding between the SOC/CRAC polypeptide and one or more binding targets is detected by any convenient method available to the user. For cell free binding type assays, a separation step is often used to separate bound from unbound components. The separation step may be accomplished in a variety of ways. Conveniently, at least one of the components is immobilized on a solid substrate, from which the unbound components may be easily separated. The solid substrate can be made of a wide variety of materials and in a wide variety of shapes, e.g., microtiter plate, microbead, dipstick, resin particle, etc. The substrate preferably is chosen to maximum signal to noise ratios, primarily to minimize background binding, as well as for ease of separation and cost.

Separation may be effected for example, by removing a bead or dipstick from a reservoir, emptying or diluting a reservoir such as a microtiter plate well, rinsing a bead, particle, chromatographic column or filter with a wash solution or solvent. The separation step preferably includes multiple rinses or washes. For example, when the solid substrate is a microtiter plate, the wells may be washed several times with a washing solution, which typically includes those components of the incubation mixture that do not participate in specific bindings such as salts, buffer, detergent, non-specific protein, etc. Where the solid substrate is a magnetic bead, the beads may be washed one or more times with a washing solution and isolated using a magnet.

Detection may be effected in any convenient way for cell-based assays such as two- or three-hybrid screens. The transcript resulting from a reporter gene transcription assay of SOC/CRAC polypeptide interacting with a target molecule typically encodes a directly or indirectly detectable product, e.g., β -galactosidase activity, luciferase activity, and the like. For cell-free binding assays, one of the components usually comprises, or is coupled to, a detectable label. A wide variety of labels can be used, such as those that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc.) or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseshoe peroxidase, etc.). The label may be bound to a SOC/CRAC binding partner, or incorporated into the structure of the binding partner.

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly

detected through optical or electron density, radioactive emissions, nonradiative energy transfers, etc. or indirectly detected with antibody conjugates, strepavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art.

Of particular importance in any of the foregoing assays and binding studies is the use of a specific sequence motif identified in the SOC-2/CRAC-1 polypeptide sequence as a kinase catalytic domain. According to the invention, amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24) (or a fragment thereof), show a localized homology with the catalytic domains of eukaryotic elongation factor-2 kinase (eEF-2 kinase, GenBank Acc. no. U93850) and *Dictyostelium* myocin heavy chain kinase A (MHCK A, GenBank Acc. no. U16856), as disclosed in Ryazanov AG, et al., *Proc Natl Acad Sci U S A*, 1997, 94(10):4884-4889. Therefore, according to the invention, a method for identifying agents useful in the modulation of SOC/CRAC polypeptide kinase activity is provided. The method involves contacting a SOC/CRAC polypeptide with kinase activity, that includes, for example, amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24) with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity; detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and comparing the kinase activity in the previous step with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC kinase activity. Other controls for kinase activity can also be performed at the same time, for example, by utilizing eEF-2 kinase and/or *Dictyostelium* MHC Kinase A, in a similar manner to the SOC/CRAC member. Methods for performing such kinase activity assays are well known in the art.

The invention thus provides SOC/CRAC-specific binding agents, methods of identifying and making such agents, and their use in diagnosis, therapy and pharmaceutical development. For example, SOC/CRAC-specific agents are useful in a variety of diagnostic and therapeutic applications, especially where disease or disease prognosis is associated with altered SOC/CRAC and SOC/CRAC calcium channel fluxing characteristics. Novel SOC/CRAC-specific binding agents include SOC/CRAC-specific antibodies and other natural intracellular and extracellular binding agents identified with assays such as two hybrid screens, and non-natural intracellular and extracellular binding agents identified in screens of chemical libraries and the like.

In general, the specificity of SOC/CRAC binding to a specific molecule is determined by binding equilibrium constants. Targets which are capable of selectively binding a SOC/CRAC polypeptide preferably have binding equilibrium constants of at least about 10^7 M^{-1} , more preferably at least about 10^8 M^{-1} , and most preferably at least about 10^9 M^{-1} . The wide variety of cell based and cell free assays may be used to demonstrate SOC/CRAC-specific binding. Cell based assays include one, two and three hybrid screens, assays in which SOC/CRAC-mediated transcription is inhibited or increased, etc. Cell free assays include SOC/CRAC-protein binding assays, immunoassays, etc. Other assays useful for screening agents which bind SOC/CRAC polypeptides include fluorescence resonance energy transfer (FRET), and electrophoretic mobility shift analysis (EMSA).

Various techniques may be employed for introducing nucleic acids of the invention into cells, depending on whether the nucleic acids are introduced *in vitro* or *in vivo* in a host. Such techniques include transfection of nucleic acid- $CaPO_4$ precipitates, transfection of nucleic acids associated with DEAE, transfection with a retrovirus including the nucleic acid of interest, liposome mediated transfection, and the like. For certain uses, it is preferred to target the nucleic acid to particular cells. In such instances, a vehicle used for delivering a nucleic acid of the invention into a cell (e.g., a retrovirus, or other virus; a liposome) can have a targeting molecule attached thereto. For example, a molecule such as an antibody specific for a surface membrane protein on the target cell or a ligand for a receptor on the target cell can be bound to or incorporated within the nucleic acid delivery vehicle. For example, where liposomes are employed to deliver the nucleic acids of the invention, proteins which bind to a surface membrane protein associated with endocytosis may be incorporated into the liposome formulation for targeting and/or to facilitate uptake. Such proteins include capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half life, and the like. Polymeric delivery systems also have been used successfully to deliver nucleic acids into cells, as is known by those skilled in the art. Such systems even permit oral delivery of nucleic acids.

Other delivery systems can include time-release, delayed release or sustained release delivery systems. Such systems can avoid repeated administrations of the anti-inflammatory agent, increasing convenience to the subject and the physician. Many types of release delivery systems are available and known to those of ordinary skill in the art. They include polymer base systems such as poly(lactide-glycolide), copolyoxalates, polycaprolactones,

polyesteramides, polyorthoesters, polyhydroxybutyric acid, and polyanhydrides. Microcapsules of the foregoing polymers containing drugs are described in, for example, U.S. Patent 5,075,109. Delivery systems also include non-polymer systems that are: lipids including sterols such as cholesterol, cholesterol esters and fatty acids or neutral fats such as mono- di- and tri-glycerides; hydrogel release systems; sytastic systems; peptide based systems; wax coatings; compressed tablets using conventional binders and excipients; partially fused implants; and the like. Specific examples include, but are not limited to: (a) erosional systems in which an agent of the invention is contained in a form within a matrix such as those described in U.S. Patent Nos. 4,452,775, 4,675,189, and 5,736,152, and (b) diffusional systems in which an active component permeates at a controlled rate from a polymer such as described in U.S. Patent Nos. 3,854,480, 5,133,974 and 5,407,686. In addition, pump-based hardware delivery systems can be used, some of which are adapted for implantation.

Use of a long-term sustained release implant may be particularly suitable for treatment of chronic conditions. Long-term release, as used herein, means that the implant is constructed and arranged to deliver therapeutic levels of the active ingredient for at least 30 days, and preferably 60 days. Long-term sustained release implants are well-known to those of ordinary skill in the art and include some of the release systems described above.

The invention also contemplates gene therapy. The procedure for performing *ex vivo* gene therapy is outlined in U.S. Patent 5,399,346 and in exhibits submitted in the file history of that patent, all of which are publicly available documents. In general, it involves introduction *in vitro* of a functional copy of a gene into a cell(s) of a subject which contains a defective copy of the gene, and returning the genetically engineered cell(s) to the subject. The functional copy of the gene is under operable control of regulatory elements which permit expression of the gene in the genetically engineered cell(s). Numerous transfection and transduction techniques as well as appropriate expression vectors are well known to those of ordinary skill in the art, some of which are described in PCT application WO95/00654. *In vivo* gene therapy using vectors such as adenovirus, retroviruses, herpes virus, and targeted liposomes also is contemplated according to the invention. See, e.g., U.S. Patent Nos. 5,670,488, entitled "Adenovirus Vector for Gene Therapy", issued to Gregory et al., and 5,672,344, entitled "Viral-Mediated Gene Transfer System", issued to Kelley et al.

The invention will be more fully understood by reference to the following examples. These examples, however, are merely intended to illustrate the embodiments of the invention and are not to be construed to limit the scope of the invention.

Examples

5 As an initial approach to identifying SOC/CRAC channels, we considered publicly available data and hypothesized that the following characteristics are likely to be exhibited by SOC/CRAC calcium channels: i) SOC/CRAC calcium channels would be integral membrane proteins related (probably distantly) to one of the known calcium channel families (e.g. voltage gated, ligand gated, Trp), and therefore should have a pore region formed by a
10 tetramer of 6-7 transmembrane (TM) regions; ii) high calcium selectivity was likely to come at the price of complexity, and therefore these were likely to be large proteins; iii) the high calcium selectivity of this type of channel was likely to be useful and, therefore, highly conserved; and iv) these channels should be expressed in one or more types of lymphocytes, since ICRAC is best defined in those cell types. Since the full genome of the nematode *C. elegans* is nearing completion, and IP3-dependent calcium signals have recently been shown
15 to be required for one or more aspects of *C. elegans* development, we took the set of proteins encoded by this genome (at the time this search was initiated WORMPEP14 was the available predicted protein set) and began searching for proteins which fit the criteria above. This search began by proceeding in alphabetical order through WORMPEP14 and arbitrarily
20 excluding all proteins below approximately 1000 amino acids in size, followed by focusing on remaining proteins with clear TM spanning regions similar to those of other calcium channels. We stopped this screen on encountering a protein designated C05C12.3, a predicted protein of 1816 amino acids (SEQ ID NO:13). C05C12.3 was notable because its central pore region had some sequence similarity to but was clearly distinct from members of the Trp
25 family of calcium channels, and the hydrophobicity plot of this region showed a characteristically wide spacing between the fifth and sixth TM regions for the amino acid residues which are thought to line the channel pore region and mediate the calcium selectivity of the channels. In addition, it lacked any ankyrin repeats in the region amino-terminal to its pore region, further distinguishing it from other Trp family proteins.

30 We then used C05C12.3 for BLAST alignment screening of the rest of the *C. elegans* genome and also mammalian databases for homologous proteins, revealing two other *C. elegans* homologues (SEQ ID NO:14 and SEQ ID NO:15), and also a recently cloned mammalian protein named melastatin-1 (MLSN-1/SOC-1, SEQ ID NOs:9 and 10, and

GenBank Acc. No. AF071787). Using these sequences, we subsequently performed an exhaustive screening of publicly accessible EST databases in search of lymphocyte homologues, but were unsuccessful in detecting any homologous transcripts in any lymphocyte lines. Since MLSN-1 (SEQ ID NOs:9 and 10) was expressed exclusively in melanocytes and retina by Northern blot hybridization and by EST database searching, there was no evidence that this type of channel was expressed in the type of cell in which ICRAC-like currents were best defined. Subsequent BLAST searches picked up mouse EST sequence AI098310 (SEQ ID NO:22) from a monocyte cell line. The I.M.A.G.E. consortium clone containing the above-identified EST was then purchased from ATCC (clone ID. 1312756, Manassas, VA) and was further characterized. Using other portions of this sequence in EST searches, we subsequently picked up similar sequences in human B-cells (SEQ ID NOs:20 and 21), and other cell types as well (SEQ ID NOs: 11, 12, 16, 17, 18, and 19). Most of these sequences were subsequently identified to be part of the 3'-UTR or of the carboxy terminal region of the proteins, which are not readily identifiable as Trp channels, providing an explanation for the art's inability to detect any type of Trp related transcripts in lymphocytes. Partial sequences from the 5' and/or 3' ends of the above identified clones were then used to screen leukocyte and kidney cDNA libraries to extend the original sequences more toward the 5' and/or 3' ends.

In view of the foregoing, it was concluded that channels of this type were expressed in many types of lymphocytes, and therefore were members of a new family of SOC/CRAC calcium channels.

Experimental Procedures Screening of the cDNA libraries

Leukocyte and kidney cDNA libraries from Life Technologies (Gaithersburg, MD) were screened using the Gene Trapper II methodology (Life Technologies) according to manufacturer's recommendation, using the inserts of I.M.A.G.E. clone ID nos. 1312756 and 1076485 from ATCC (Manassas, VA), under stringent hybridization conditions. Using standard methodology (*Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York), individual cDNA clones were subjected to 3-4 rounds of amplification and purification under the same hybridization conditions.

After excision from the vector and subcloning of inserts into the plasmid forms, several clones were sequenced by the Beth Israel Deaconess Medical Center's Automated

Sequencing Facility. Molecular biological techniques such as restriction enzyme treatment, subcloning, DNA extraction, bacterial culture and purification of DNA fragments were performed according to methods well known in the art. Computer analyses of protein and DNA sequences was done using "Assemblylign" (Oxford Molecular, Cambell, CA). Multiple alignments of the SOC/CRAC family members were produced using the CLUSTAL facility of the MacVector program. Restriction endonucleases, expression vectors, and modifying enzymes were purchased from commercial sources (Gibco-BRL). Sequencing vectors for DNA were purchased from Stratagene (La Jolla, CA).

Once the first members of what appeared to be a novel family of calcium channel receptors were identified and characterized, additional BLAST alignments were performed with the newly characterized nucleic acid sequences. An initial match was with genomic DNA fragment NH0332L11 (Genbank Acc. No. AC005538). Using this genomic sequence, promoters were designed and a number of cDNA libraries was surveyed by PCR. A prostate specific message was identified and characterized, leading to the isolation and characterization of SOC-4/CRAC-3 (SEQ ID NOs: 31 and 32).

Functional Assays

Transient Expression of SOC/CRAC

In our initial transient expression experiments, we expressed or expect to express a SOC/CRAC molecule transiently in RBL-2H3 mast cells, Jurkat T cells, and A20 B-lymphocytes using both electroporation and vaccinia virus-driven expression, and measured the calcium influx produced by depletion of intracellular calcium stores with thapsigargin. Each of the foregoing techniques is well known to those of ordinary skill in the art and can be performed using various methods (see, e.g., Current Methods in Molecular Biology, eds. Ausubal, F.M., et al. 1987, Green Publishers and Wiley Interscience, N.Y., N.Y.). Exemplary methods are described herein.

Depletion of intracellular calcium stores is accomplished by treating the cells with 1 micromolar thapsigargin; alternative agents which function to deplete intracellular stores are described in by Putney, J.W. Jr., in Capacitative Calcium Entry, R.G. Landes Co. and Chapman & Hall, 1997 and include, for example, ionomycin, cyclopiazonic acid, and DBHQ.

Calcium influx is determined by measuring cytoplasmic calcium as indicated using the fura-2 fluorescent calcium indicator (see, e.g., G. Grynkiewicz, M. Poenie, R. Y. Tsien, A new generation of Ca²⁺ indicators with greatly improved fluorescence properties, J. Biol

Chem 260, 3440-50 (1985), and M. Poenie, R. Tsien, Fura-2: a powerful new tool for measuring and imaging $[Ca^{2+}]_i$ in single cells, Prog Clin Biol Res 210, 53-6 (1986)).

Patch Clamp Analysis and Determining Selectivity of SOC/CRAC

Patch clamp analysis of cells injected with SOC/CRAC cRNA is performed by using the general patch technique as described in Neher, E., "Ion channels for communication between and within cells", Science, 1992; 256:498-502. Specific techniques for applying the patch clamp analysis to RBL cells are described in Hoth, M., and Penner, R., "Depletion of intracellular calcium stores activates a calcium current in mast cells", Nature, 1992; 355:3535-355. Additional protocols for applying the patch clamp technique to other cell types are described in Putney, J.W. Jr., in Capacitative Calcium Entry, R.G. Landes Co. and Chapman & Hall, 1997

An exemplary protocol for patch clamp analysis of SOC/CRAC molecule expressed in RBL-2H3 mast cells using a recombinant vaccinia virus is as follows. The currents elicited by store depletion are determined using the whole cell configuration (Neher, E., Science, 1992; 256:498-502). Currents in SOC/CRAC expressing cells are compared to currents in control cells expressing an irrelevant protein or a classic Trp family calcium channel known as VR1 (M. J. Caterina, et al., The capsaicin receptor: a heat-activated ion channel in the pain pathway [see comments], Nature 389, 816-24 (1997)) in order to assess the contribution of SOC/CRAC expression. In addition, the magnitude of whole cell currents in the presence of extracellular calcium (10 mM), barium (10 mM), or magnesium (10 mM) are compared to determine the relative permeability of the channels to each of these ions (Hoth, M., and Penner, R., Nature, 1992; 355:3535-355) and, thereby, determine the ionic selectivity.

Pharmacologic Behavior of SOC/CRAC

For analysis of the pharmacologic behavior of a SOC/CRAC molecule, a SOC/CRAC molecule is expressed in RBL-2H3 mast cells using a recombinant vaccinia virus, and the degree of calcium influx elicited by store depletion is monitored using a bulk spectrofluorimeter or a fluorescence microscope and the calcium sensitive dye fura-2 (G. Grynkiewicz, M. Poenie, R. Y. Tsien, A new generation of Ca^{2+} indicators with greatly improved fluorescence properties, J Biol Chem 260, 3440-50 (1985) and M. Poenie, R. Tsien, Fura-2: a powerful new tool for measuring and imaging $[Ca^{2+}]_i$ in single cells, Prog Clin Biol Res 210, 53-6 (1986)). The level of cytoplasmic calcium in SOC/CRAC expressing cells is compared to the level achieved in control cells expressing an irrelevant protein or a classic Trp. family calcium channels known as VR1 (M. J. Caterina, et al., The

capsaicin receptor: a heat-activated ion channel in the pain pathway [see comments], Nature 389, 816-24 (1997)). These cells then are pre-incubated with the desired pharmacologic reagent, and again the response to store depletion is monitored. Comparison of the effect of depleting stores in SOC/CRAC expressing cells relative to controls in the presence or absence of the pharmacologic reagent is used to assess the ability of that reagent to modulate SOC/CRAC activity. Sphingosine is an exemplary molecule that can be used as pharmacologic reagents for pharmacologic characterization of SOC/CRAC calcium channels. See, e.g., Mathes, C., et al., Calcium release activated calcium current as a direct target for sphingosine, J Biol Chem 273(39):25020-25030 (1998). Other non-specific calcium channel inhibitors that can be used for this purpose include SKR96365 (Calbiochem) and Lanthanum.

Bulk Calcium Assays

Bulk calcium assays can be performed in a PTI Deltascan bulk spectrofluorometer using fura-2 as described in Scharenberg AM, et al., EMBO J, 1995, 14(14):3385-94.

Gene Targeting

The method (and reagents) described by Buerstedde JM et al, (Cell, 1991, Oct 4;67(1):179-88), was used to generate "knockouts" in cells. Briefly, part of the chicken SOC-2/CRAC-1 genomic sequence coding for the transmembrane region was cloned utilizing the human sequence as the probe in a chicken library screen. Chicken SOC-2/CRAC-1 clones were isolated and characterized using standard methodology. The putative exon and domain arrangement of the chicken SOC-2/CRAC-1, is depicted in Figure 1. The exons coding for TM5 (pore region) and TM6, were replaced with promoter/antibiotic cassettes (see Figure1). These targeting vectors were then used to target (and replace) the endogenous gene in DT-40 cells (chicken B lymphocyte cells).

Results

Example 1: Transient Expression of SOC/CRAC

In the above-identified cell lines and using both of the foregoing expression techniques, SOC/CRAC expression enhances thapsigargin-dependent influx. In addition, SOC/CRAC expression also enhances the amount of intracellular calcium stores. That this effect is likely due to SOC/CRAC acting as a plasma membrane calcium channel can be confirmed by producing an in-frame carboxy-terminal translational fusion with green fluorescent protein followed by confocal microscopy, revealing that SOC/CRAC is expressed predominantly as a plasma membrane calcium channel.

Example 2: Patch Clamp Analysis

The biophysical characteristics of SOC/CRAC enhanced currents when expressed in *Xenopus* oocytes are determined. SOC/CRAC cRNA injection is able to enhance thapsigargin-dependent whole cell currents. In addition, SOC/CRAC does not alter the reversal potential of these currents and the determination of the P_{Ca}/P_{Na} ratio shows that SOC/CRAC channels are highly calcium selective.

Example 3: *Pharmacologic Behavior of SOC/CRAC*

The pharmacologic behavior of SOC/CRAC is evaluated as described above. SOC/CRAC-enhanced influx is inhibited by sphingosine in a manner that is substantially the same as that of endogenous thapsigargin-dependent calcium influx.

Example 4: *Gene targeting*

Transfection of DT-40 cells with the foregoing targeting vectors, selection for antibiotic resistance, and screening, is collectively referred to, herein, as a round of targeting. For the first round of targeting SOC-2/CRAC-1, 18/24 clones with homologous recombination of the targeting construct into one of the endogenous SOC-2/CRAC-1 alleles were obtained. On the second round of targeting (in order to target the second allele and therefore generate a homozygous SOC-2/CRAC-1 mutant cell), 0/48 clones were obtained. These results indicate that a "null" SOC-2/CRAC-1 mutation is detrimental to DT-40 cells, and that SOC-2/CRAC-1 is required for cell viability.

Table I. Nucleotide Sequences with homologies to SOC/CRAC nucleic acids

Sequences with SEQ ID NOs and GenBank accession numbers:
SEQ ID NO:9, AB001535, AI226731, H18835, AA419592, AA261842, AA419407, AA592910, D86107, AI098310, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853.

Table II. Amino Acid Sequences with homologies to SOC/CRAC polypeptides

Sequences with SEQ ID NOs and GenBank accession numbers:
SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572.

All references, patents, and patent documents disclosed herein are incorporated by reference herein in their entirety.

What is claimed is presented below and is followed by a Sequence Listing. We claim:

Claims

1. An isolated nucleic acid molecule, comprising:

(a) nucleic acid molecules which hybridize under stringent conditions to a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;

(b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;

(c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and

(d) complements of (a), (b) or (c).

2. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:1.

3. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:27.

4. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:29.

5. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:31.

6. An isolated nucleic acid molecule selected from the group consisting of

(a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:29, and SEQ ID NO:31,

(b) complements of (a),

provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to any sequence selected from a sequence group consisting of

(1) sequences having the SEQ. ID NOS. or GenBank accession numbers of Table I,

(2) complements of (1), and

(3) fragments of (1) and (2).

7. The isolated nucleic acid molecule of claim 6, wherein the sequence of contiguous nucleotides is selected from the group consisting of:

- (1) at least two contiguous nucleotides nonidentical to the sequence group,
- (2) at least three contiguous nucleotides nonidentical to the sequence group,
- 5 (3) at least four contiguous nucleotides nonidentical to the sequence group,
- (4) at least five contiguous nucleotides nonidentical to the sequence group,
- (5) at least six contiguous nucleotides nonidentical to the sequence group,
- (6) at least seven contiguous nucleotides nonidentical to the sequence group.

10 8. The isolated nucleic acid molecule of claim 6, wherein the unique fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, and 200 nucleotides.

15 9. The isolated nucleic acid molecule of claim 6, wherein the molecule encodes a polypeptide which is immunogenic.

10. An expression vector comprising the isolated nucleic acid molecule of claims 1, 2, 3, 4, 5, 6, 7, 8, or 9 operably linked to a promoter.

11. A host cell transformed or transfected with the expression vector of claim 10.

20 12. An isolated polypeptide encoded by the isolated nucleic acid molecule according to anyone of claims 1 or 6, wherein the polypeptide comprises a SOC/CRAC polypeptide or a unique fragment thereof.

13. The isolated polypeptide of claim 12, wherein the isolated polypeptide is encoded by the isolated nucleic acid molecule of claim 2, 3, 4, or 5.

25 14. The isolated polypeptide of claim 13, wherein the isolated polypeptide comprises a polypeptide having the sequence of amino acids selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32.

15. An isolated polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3, 4, or 5, wherein the polypeptide, or unique fragment thereof is immunogenic.

16. An isolated binding polypeptide which binds selectively to a polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3, 4, or 5.

5 17. The isolated binding polypeptide of claim 16, wherein the isolated binding polypeptide binds to a polypeptide having the sequence of amino acids selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32.

10 18. The isolated binding polypeptide of claim 17, wherein the isolated binding polypeptide is an antibody or an antibody fragment selected from the group consisting of a Fab fragment, a F(ab)₂ fragment or a fragment including a CDR3 region selective for the polypeptide.

15 19. An isolated polypeptide, comprising a unique fragment of the polypeptide of claim 12 of sufficient length to represent a sequence unique within the human genome, provided that the fragment excludes a sequence of contiguous amino acids identified in Table II.

20. A method for isolating a SOC/CRAC molecule having SOC/CRAC calcium channel activity, comprising:

20 a) contacting a binding molecule that is a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing one or more SOC/CRAC molecules, under conditions sufficient to form a complex of the SOC/CRAC nucleic acid or the SOC/CRAC binding polypeptide and the SOC/CRAC molecule;

b) detecting the presence of the complex;

c) isolating the SOC/CRAC molecule from the complex; and

25 d) determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity.

21. The method of claim 20, wherein the binding molecule is a SOC/CRAC nucleic acid.

22. The method of claim 20, wherein the binding molecule is a SOC/CRAC binding polypeptide.

23. The method of claim 21, wherein the SOC/CRAC nucleic acid comprises at least 14 nucleotides from any contiguous portion of a sequence of nucleotides selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31.

5 24. A method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity, comprising:

a) contacting a SOC/CRAC polypeptide with a candidate agent suspected of modulating SOC/CRAC calcium channel activity, under conditions sufficient to allow the SOC/CRAC polypeptide to interact selectively with the candidate agent;

10 b) detecting a Ca^{2+} concentration associated with SOC/CRAC calcium channel activity of the SOC/CRAC polypeptide in the presence of the candidate agent; and

c) comparing the Ca^{2+} concentration of step (b) with a control Ca^{2+} concentration of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates SOC/CRAC calcium channel activity.

15 25. A method for determining the level of SOC/CRAC expression in a subject, comprising:

a) measuring the expression of SOC/CRAC in a test sample obtained from the subject, and

20 b) comparing the measured expression of SOC/CRAC in the test sample to the expression of the SOC/CRAC polypeptide in a control to determine the level of SOC/CRAC expression in the subject.

25 26. The method of claim 25, wherein the expression of SOC/CRAC in (b) is SOC/CRAC mRNA expression.

27. The method of claim 25, wherein the expression of SOC/CRAC in (b) is SOC/CRAC polypeptide expression.

28. The method of claim 25, wherein the test sample is tissue.

29. The method of claim 25, wherein the test sample is a biological fluid.

30. The method of claim 26, wherein SOC/CRAC mRNA expression is measured using the Polymerase Chain Reaction (PCR).

31. The method of claim 26, wherein SOC/CRAC mRNA expression is measured using a method selected from the group consisting of northern blotting, monoclonal antisera to SOC/CRAC and polyclonal antisera to SOC/CRAC.

32. A kit, comprising a package containing:

an agent that selectively binds to the isolated nucleic acid of claim 1 or an expression product thereof, and

a control for comparing to a measured value of binding of said agent to said isolated nucleic acid of claim 1 or expression product thereof.

33. The kit of claim 32, wherein the control comprises an epitope of the expression product of the nucleic acid of claim 1.

34. A pharmaceutical composition comprising:

a pharmaceutically effective amount of an agent comprising of an isolated nucleic acid molecule of claim 1 or an expression product thereof, and

a pharmaceutically acceptable carrier.

35. The pharmaceutical composition of claim 34, wherein the agent is an expression product of the isolated nucleic acid molecule of claim 1.

36. A method for identifying agents useful in the modulation of a SOC/CRAC polypeptide kinase activity, comprising:

a) contacting a SOC/CRAC polypeptide with kinase activity with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity;

b) detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and

c) comparing the kinase activity of step (b) with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates SOC/CRAC kinase activity.

37. The method of claim 36, wherein the SOC/CRAC polypeptide comprises amino acids 999-1180 of the sequence represented as SEQ ID NO:24, or a fragment thereof that retains the kinase activity.

1/1

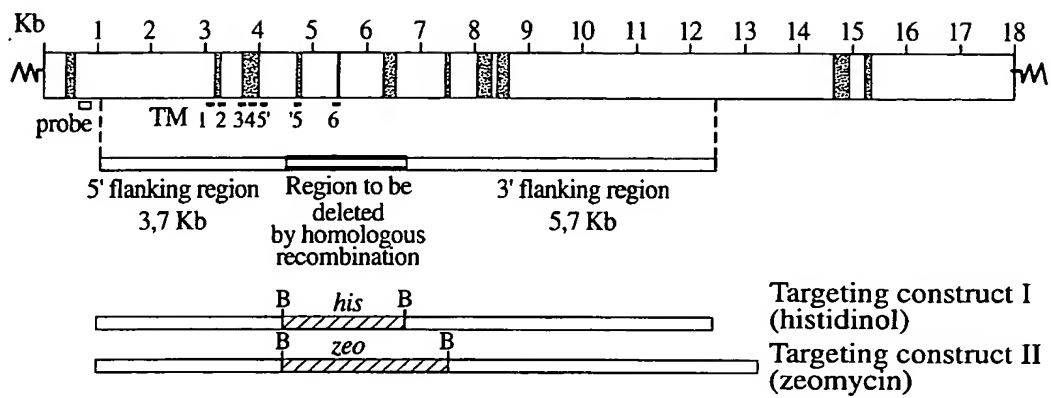


Fig. 1

-1-

SEQUENCE LISTING

<110> Beth Israel Deaconess Medical Center, Inc.
Scharenberg, Andrew

<120> CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY

<130> B0662/7026WO/ERP/KA

<150> U.S. 60/114,220

<151> 1998-12-30

<150> U.S. 60/120,018

<151> 1999-01-29

<150> U.S. 60/140,415

<151> 1999-06-22

<160> 32

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 1212

<212> DNA

<213> Homo Sapiens

<400> 1

```
gcacgaggca aattttttgt tagtacacca tctcagccaa gttgcaaaaag ccacttggaa      60
actggaacca aagatcaaga aactgtttgc tctaaagcta cagaaggaga taatacagaa      120
tttggagcat ttgtaggaca cagagatagc atggatttac agaggtttaa agaaacatca      180
aacaagataa aaatactatc caataacaat acttctgaaa acactttgaa acgagtgagt      240
tctcttgctg gattttactga ctgtcacaga acttccattc ctgttcattc aaaacaagaa      300
aaaatcagta gaaggccatc taccgaagac actcatgaag tagattccaa agcagcttta      360
ataccggttt gtagatttca actaaacaga tatatatatt taaatacatt aaactttttt      420
agataagatc tacaaagtgg tgatatattg gactatatca aaaattcaaa aaaatttttc      480
ttaagaaaac tgacttttagc atagtagcag ttacagaaaa gtttcttaca gtgaatagtc      540
aggaatttta aagaaaaatt tatgcagaat aaaggcagga atctcttttt gtttgaattg      600
aagctaatta tatgaactca tttccagcta actgcgataa tgattgattt tgcaaattcc      660
ctttaaaagc acacactgac aagacaaaaa gctcaggaaa aggcagaaaa attactcctt      720
tataatcaag tattatatat aagtcagtgc tcataatttt gctcaagaaa atattgactt      780
acattcatat atatctgttc tggcatagag agattatggt gttaaaaatca tgttattgaa      840
aaaagttatt tcagtgggga aagaggttag ttaacaaaga gattcacagt aacaaatcct      900
cctttctgga gggactcttc ctgaccctga gctgcacaac ttgcaacaa attaaagcct      960
aaccgaagat gacctcaciaa tggcaattta gaactcatgg gagtcaactt acataaacgg      1020
tatttgattt ctgataagat agtggaatta ttgggttatag atgacaaaat aagtatgttt      1080
aaagtgatga tggacataaa aaagttttta atataaaaca tgagaaaaga aggagatact      1140
attcaaaaag actggcaaatt ttgaaaaact agaaataaaa aaaaaaaaaa aaaatgagcg      1200
gccgcaagct tt                                     1212
```

<210> 2

<211> 141

<212> PRT

<213> Homo Sapiens

<400> 2

```
Ala Arg Gly Lys Phe Phe Val Ser Thr Pro Ser Gln Pro Ser Cys Lys
  1          5          10          15
Ser His Leu Glu Thr Gly Thr Lys Asp Gln Glu Thr Val Cys Ser Lys
  20          25          30
```


-2-

Ala	Thr	Glu	Gly	Asp	Asn	Thr	Glu	Phe	Gly	Ala	Phe	Val	Gly	His	Arg
	35						40					45			
Asp	Ser	Met	Asp	Leu	Gln	Arg	Phe	Lys	Glu	Thr	Ser	Asn	Lys	Ile	Lys
	50					55					60				
Ile	Leu	Ser	Asn	Asn	Asn	Thr	Ser	Glu	Asn	Thr	Leu	Lys	Arg	Val	Ser
	65					70				75					80
Ser	Leu	Ala	Gly	Phe	Thr	Asp	Cys	His	Arg	Thr	Ser	Ile	Pro	Val	His
			85						90					95	
Ser	Lys	Gln	Glu	Lys	Ile	Ser	Arg	Arg	Pro	Ser	Thr	Glu	Asp	Thr	His
			100					105					110		
Glu	Val	Asp	Ser	Lys	Ala	Ala	Leu	Ile	Pro	Val	Cys	Arg	Phe	Gln	Leu
		115					120					125			
Asn	Arg	Tyr	Ile	Leu	Leu	Asn	Thr	Leu	Asn	Phe	Phe	Arg			
	130					135					140				

<210> 3
 <211> 739
 <212> DNA
 <213> Homo Sapiens

<220>
 <221> unsure
 <222> (5)...(5)
 <223> UNKNOWN

<221> unsure
 <222> (21)...(22)
 <223> UNKNOWN

<221> unsure
 <222> (29)...(29)
 <223> UNKNOWN

<400> 3	
tcgantaggg	gtcttccacc nncatactng gatgatgggt ggtgaagtct atgcatacga 60
aattgatgtg	tgtgcaaacg attctgttat ccctcaaatac tgtggctcctg ggacgtgggt 120
gactccattt	cttcaagcag tctacctctt tgwacagtat atcattatgg ttaattcttct 180
tattgcattt	ytcaacaatg tgtattttaca agtgaaggca atttccaata ttgyatggaa 240
gtaccagcgt	tatcatttta ttatggctta tcatgagaaa ccagttctgc ctccctccact 300
tatcattctt	agccatatag tttctctgtt ttgctgcata tgtaagagaa gaaagaaaga 360
taagacttcc	gatggaccaa aacttttctt aacagaagaa gatcaaaaaga aacttcatga 420
ttttgaagag	cagtgtgttg aaatgtattt caatgaaaaa gatgacaaat ttcattctgg 480
gagtgaagag	agaattcgtg tcacttttga aagagtggaa cagatgtgca ttcagattaa 540
agaagttgga	gatccgtgtc aactacataa aaagatcatt acaatcatga gattctcaaa 600
ttggccattt	gcaagatctt tcagccctga cggtagatac attaaaaaca ctactggcc 660
aaaagcgtcg	gaagctagca aagttcataa tgaaatcaca cgagaactga gcatttccaa 720
acacttggct	caaaacctt

<210> 4
 <211> 235
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> UNSURE
 <222> (41)...(41)
 <223> UNKNOWN

<221> UNSURE
 <222> (54)...(54)

-3-

<223> UNKNOWN

<221> UNSURE

<222> (68)...(68)

<223> UNKNOWN

<400> 4

Met	Met	Val	Gly	Glu	Val	Tyr	Ala	Tyr	Glu	Ile	Asp	Val	Cys	Ala	Asn
1				5					10					15	
Asp	Ser	Val	Ile	Pro	Gln	Ile	Cys	Gly	Pro	Gly	Thr	Trp	Leu	Thr	Pro
			20					25					30		
Phe	Leu	Gln	Ala	Val	Tyr	Leu	Phe	Xaa	Gln	Tyr	Ile	Ile	Met	Val	Asn
		35					40					45			
Leu	Leu	Ile	Ala	Phe	Xaa	Asn	Asn	Val	Tyr	Leu	Gln	Val	Lys	Ala	Ile
		50				55					60				
Ser	Asn	Ile	Xaa	Trp	Lys	Tyr	Gln	Arg	Tyr	His	Phe	Ile	Met	Ala	Tyr
65					70					75					80
His	Glu	Lys	Pro	Val	Leu	Pro	Pro	Pro	Leu	Ile	Ile	Leu	Ser	His	Ile
				85					90					95	
Val	Ser	Leu	Phe	Cys	Cys	Ile	Cys	Lys	Arg	Arg	Lys	Lys	Asp	Lys	Thr
			100					105					110		
Ser	Asp	Gly	Pro	Lys	Leu	Phe	Leu	Thr	Glu	Glu	Asp	Gln	Lys	Lys	Leu
		115					120					125			
His	Asp	Phe	Glu	Glu	Gln	Cys	Val	Glu	Met	Tyr	Phe	Asn	Glu	Lys	Asp
		130				135					140				
Asp	Lys	Phe	His	Ser	Gly	Ser	Glu	Glu	Arg	Ile	Arg	Val	Thr	Phe	Glu
145					150					155					160
Arg	Val	Glu	Gln	Met	Cys	Ile	Gln	Ile	Lys	Glu	Val	Gly	Asp	Pro	Cys
				165					170					175	
Gln	Leu	His	Lys	Lys	Ile	Ile	Thr	Ile	Ile	Arg	Phe	Ser	Asn	Trp	Pro
			180					185					190		
Phe	Ala	Arg	Ser	Phe	Ser	Pro	Asp	Gly	Arg	Tyr	Ile	Lys	Asn	Thr	His
		195				200						205			
Trp	Pro	Lys	Ala	Ser	Glu	Ala	Ser	Lys	Val	His	Asn	Glu	Ile	Thr	Arg
		210				215					220				
Glu	Leu	Ser	Ile	Ser	Lys	His	Leu	Ala	Gln	Asn					
225					230					235					

<210> 5

<211> 1579

<212> DNA

<213> Homo Sapiens

<220>

<221> unsure

<222> (368)...(368)

<223> g or c

<221> unsure

<222> (372)...(372)

<223> g or c

<221> unsure

<222> (374)...(374)

<223> g or a

<221> unsure

<222> (375)...(375)

<223> g or c

-4-

<221> unsure
 <222> (387) ... (387)

<221> unsure
 <222> (482) ... (482)

<400> 5

acgtgcgctg	caggtaccgg	tccggaattc	ccgggtcgac	ccacgcgtcc	ggcatgggtgt	60
tgtaaatata	cttagctcct	ctcttcctca	aggtgatctt	gaaagtaata	atccttttca	120
ttgtaatat	ttaatgaaag	atgacaaaga	tccccagtgt	aatatatattg	gtcaagactt	180
acctgcagta	ccccagagaa	aagaatttaa	ttttccagag	gctgggtcct	cttctgggtgc	240
cttattccca	agtgcgtgtt	cccctccaga	actgcgacag	agactacatg	gggtagaact	300
cttaaaaaata	tttaataaaaa	atcaaaaaatt	aggcagttca	tctactagca	taccacatct	360
gtcatccsca	csarscaa	tttttgntag	tacaccatct	cagccaagtt	gcaaaagcca	420
cttggaact	ggaaccaaag	atcaagaaac	tgtttgctct	aaagctacag	aaggagataa	480
tncagaattt	ggagcatttg	taggacacag	agatagcatg	gatttacaga	ggtttaaaga	540
aacatcaaac	aagataaaaa	tactatccaa	taacaatact	tctgaaaaca	ctttgaaacg	600
agtgagttct	cttgctggat	ttactgactg	tcacagaact	tccattcctg	ttcattcaaa	660
acaagaaaaa	atcagtagaa	ggccatctac	cgaagacact	catgaagtag	attccaaagc	720
agctttaata	ccggtttgta	gatttcaact	aaacagatat	atattattaa	atacattaaa	780
cttttttaga	taagatctac	aaagtgggtga	tatttgggac	tatatcaaaa	attcaaaaaa	840
atttttctta	agaaaactga	ctttagcata	gtagcagtta	cagaaaagtt	tcttacagtg	900
aatagtcagg	aattttaaag	aaaaatttat	gcagaataaa	ggcaggaatc	tctttttgtt	960
tgaattgaag	ctaattatat	gaactcattt	ccagctaact	gcgataatga	ttgattttgc	1020
aaattccctt	taaaagcaca	cactgacaag	acaaaaagct	caggaaaagg	cagaaaaaatt	1080
actcctttat	aatcaagtat	tatatataag	tcagtgtctca	taattttgct	caagaaaata	1140
ttgacttaca	ttcatatata	tctgttctgg	catagagaga	ttatgttggt	aaaatcatgt	1200
tattgaaaaa	agttatttca	gtggggaaag	aggttagtta	acaaagagat	tcacagtaac	1260
aaatcctcct	ttctggaggg	actcttcctg	accctgagct	gcacaacttt	gcaacaaatt	1320
aaagcctaac	cgaagatgac	ctcacaatgg	caatttagaa	ctcatgggag	tcaacttaca	1380
taaacggtat	ttgattttctg	ataagatagt	ggaattattg	gttatagatg	acaaaataag	1440
tatgtttaaa	gtgatgatgg	acataaaaaa	gttttaataa	taaaacatga	gaaaagaagg	1500
agatactatt	caaaaagact	ggcaaatattg	aaaaactaga	aataaaaaaa	aaaaaaaaaa	1560
atgagcgcc	gcaagcttt					1579

<210> 6
 <211> 243
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> UNSURE
 <222> (103) ... (105)
 <223> UNKNOWN

<221> UNSURE
 <222> (109) ... (109)
 <223> UNKNOWN

<221> UNSURE
 <222> (141) ... (141)
 <223> UNKNOWN

<400> 6

Val	Asn	Thr	Leu	Ser	Ser	Leu	Pro	Gln	Gly	Asp	Leu	Glu	Ser	Asn
1				5				10					15	
Asn	Pro	Phe	His	Cys	Asn	Ile	Leu	Met	Lys	Asp	Asp	Lys	Asp	Gln
			20					25				30		
Cys	Asn	Ile	Phe	Gly	Gln	Asp	Leu	Pro	Ala	Val	Pro	Gln	Arg	Lys
		35					40					45		

-5-

Phe	Asn	Phe	Pro	Glu	Ala	Gly	Ser	Ser	Ser	Gly	Ala	Leu	Phe	Pro	Ser
	50					55					60				
Ala	Val	Ser	Pro	Pro	Glu	Leu	Arg	Gln	Arg	Leu	His	Gly	Val	Glu	Leu
65					70					75					80
Leu	Lys	Ile	Phe	Asn	Lys	Asn	Gln	Lys	Leu	Gly	Ser	Ser	Ser	Thr	Ser
				85					90					95	
Ile	Pro	His	Leu	Ser	Ser	Xaa	Xaa	Xaa	Lys	Phe	Phe	Xaa	Ser	Thr	Pro
			100					105					110		
Ser	Gln	Pro	Ser	Cys	Lys	Ser	His	Leu	Glu	Thr	Gly	Thr	Lys	Asp	Gln
		115					120					125			
Glu	Thr	Val	Cys	Ser	Lys	Ala	Thr	Glu	Gly	Asp	Asn	Xaa	Glu	Phe	Gly
	130					135					140				
Ala	Phe	Val	Gly	His	Arg	Asp	Ser	Met	Asp	Leu	Gln	Arg	Phe	Lys	Glu
145					150					155					160
Thr	Ser	Asn	Lys	Ile	Lys	Ile	Leu	Ser	Asn	Asn	Asn	Thr	Ser	Glu	Asn
			165					170						175	
Thr	Leu	Lys	Arg	Val	Ser	Ser	Leu	Ala	Gly	Phe	Thr	Asp	Cys	His	Arg
			180					185					190		
Thr	Ser	Ile	Pro	Val	His	Ser	Lys	Gln	Glu	Lys	Ile	Ser	Arg	Arg	Pro
		195					200					205			
Ser	Thr	Glu	Asp	Thr	His	Glu	Val	Asp	Ser	Lys	Ala	Ala	Leu	Ile	Pro
	210					215					220				
Val	Cys	Arg	Phe	Gln	Leu	Asn	Arg	Tyr	Ile	Leu	Leu	Asn	Thr	Leu	Asn
225					230					235					240
Phe	Phe	Arg													

<210> 7
 <211> 3532
 <212> DNA
 <213> Mus Musculus

<220>
 <221> unsure
 <222> (2420)... (2420)
 <223> unknown

<221> unsure
 <222> (2434)... (2434)
 <223> unknown

<221> unsure
 <222> (2461)... (2461)
 <223> unknown

<221> unsure
 <222> (2466)... (2466)
 <223> unknown

<221> unsure
 <222> (2470)... (2470)
 <223> unknown

<400> 7
 attatggcctt atcatgaaaa accagtcctg cctcctcctc ttatcatcct cagccatata 60
 gtttcaactgt tttgctgtgt atgcaaaaaga agaaagaaag ataagacttc cgatgggcca 120
 aaacttttct taacagaaga agatcaaaaag aaactccatg attttgaaga gcagtgtggt 180
 gagatgtact ttgatgagaa agatgacaaa ttcaattctg ggagtgaaga gagaatccgg 240
 gtcacttttg aaagagtggg gcagatgagc attcagatta aagaagttgg agatcgtgtc 300
 aactacataa aaagatcatt acagtcttta gattctcaaa ttggtcatct gcaagatctc 360

-6-

tcagccctaa	cagtagatac	attgaaaaca	cttacagccc	agaaagcttc	agaagctagt	420
aaagtgcaca	atgagatcac	acgagaattg	agtatttcca	aacacttggc	tcagaatctt	480
attgatgatg	ttcctgtaag	acctttgtgg	gaagaaccta	gtgctgtaaa	cacactgagt	540
tcctctcttc	ctcaaggtga	tcgggaaagt	aataatcctt	ttctttgtaa	tatttttatg	600
aaagatgaaa	aagaccccca	atataatctg	tttggaacaag	atttgcccgt	gatacccccag	660
agaaaagaat	tcaacattcc	agaggctggg	tcctcctgtg	gtgccttatt	cccaagtgtc	720
gtttctcccc	cagaattacg	acagagacga	catggggtag	aaatgttaaa	aatatttaat	780
aaaaatcaaa	aattaggcag	ttcacctaata	agttcaccac	atatgtcctc	cccaccaacc	840
aaattttctg	tgagtacccc	atcccagcca	agttgcaaaa	gtcacttgga	atccacaacc	900
aaagatcaag	aaccattttt	ctataaagct	gcagaagggg	ataacataga	atttggagca	960
tttgtgggac	acagagatag	tatggactta	cagaggttta	aagaaacatc	aaacaaaata	1020
agagaactgt	tatctaatag	tactcctgaa	aacactctga	aacatgtggg	tgctgctgga	1080
tatagtgaat	gttgtaagac	ttctacttct	cttcactcgg	tgcaagcaga	aagctgtagt	1140
agaagagcgt	cgacggaaga	ctctccagaa	gtcgattcta	aagcagcttt	gttaccggat	1200
tggttacgag	atagaccatc	aaacagagaa	atgccatctg	aaggaggaac	attaaatggg	1260
cttgcttctc	catttaagcc	cgttttggat	aacaattact	attattcagc	tgtggaaaga	1320
aataacctga	tgaggttgct	acagagtatt	cccttcgctt	ctgtacctcc	acgaggcgag	1380
cctgtcacag	tgtaccgtct	ggaggagagt	tctcccagta	tactgaataa	cagcatgtct	1440
tcatggtctc	agctaggcct	ctgtgccaaa	attgagtttt	taagtaaaga	ggaaatggaa	1500
ggtggtttac	gaagagcagt	caaagtgtcg	tgtacctggg	cagagcacga	tatcctgaag	1560
tcagggcac	tctatatcat	taagtcattt	cttcctgagg	tgataaacac	atgggtcaagc	1620
atttataaag	aagatacggg	tctacatctc	tgtctcagag	aaatacaaca	acagagagca	1680
gcacaaaagc	tcacatttgc	ctttaatcag	atgaaaccca	aatccatacc	atattctcca	1740
aggttccttg	aagttttcct	gttgactgc	cattcagcag	ggcagtggtt	tgctgtagaa	1800
gagtgcatga	ctggtgaatt	tagaaaatac	aacaacaata	atgggtgatga	aatcattcct	1860
acaaatactc	tagaagagat	catgctagcc	tttagccact	ggacctatga	atataccaga	1920
ggggagttac	tggtacttga	cttacaagga	gtgggagaaa	acttgactga	cccatctgta	1980
ataaaagctg	aagaaaaaag	atcctgtgac	atgggtttttg	gccctgccaa	tctaggagaa	2040
gatgcaataa	aaaacttcaa	gagccaaaca	tccactgtaa	ttcttgctgt	cgaagcctta	2100
aacttcccag	atttgaagag	gaatgactac	acgcccttga	taaaattata	tttctcagg	2160
atgagtcac	agatttgaat	cttcaatctg	gaaattccac	caaagaatca	gaagcaacaa	2220
attctgttcg	tctgatgtta	tagtgctgag	tcattgggtt	ttgcctacac	ttcacaaaag	2280
tgtaactgtc	agttttcctt	tcgggggaat	tgatgatata	ggaagatgtg	tgcaaaatga	2340
gcttgctggc	cccacacata	gtctagaggt	aatgttctca	tgaaaaaacg	cctggagggtg	2400
gaggctgcag	atgccagtgn	aaagtgtctag	ctgncagaga	gtcagtgctc	tcgggctggg	2460
naaggncggn	acccttgctg	ctgagagtgg	tggttctctt	cacctgggtg	aggaccatta	2520
accaaagtca	agtcttcaga	tttgattggc	tgctcagtc	cagcccattc	agctaaggaa	2580
actaaattgc	gcagcttttt	aaatggctga	agtcttcctc	agtttgtgct	ctatgataat	2640
gatgttagct	ctcaactagg	tgtttgtggc	cacgggagaa	ctactcctta	caattttgct	2700
tcacaggcat	gttacaagc	ctgcactgaa	aaccgtttgt	cttccctctc	tccttccctc	2760
ttttccctgt	agtattgagg	atcaaaccga	gggcctcatg	aagaccattt	tctaagagac	2820
attttattta	agaatcaact	atagagtcta	tgtttatgga	tacagccagt	ttttgtttaa	2880
caaaacctga	attgtgcaaa	agggtttttt	aacatttatc	aatgttaagt	aaaagaaagc	2940
catgataaat	aagaattaac	tactgttcca	atgggtgttt	cctgtgagga	aggttacagt	3000
tgtaacagcc	tgagttgca	tacatctcca	aagattttaca	gacttagtgt	atcaaatacag	3060
agtgtcatgt	gagctctcac	attgaaaatt	ctataggaat	gtgtcaatgt	gaatttctatt	3120
tctggtactt	aagaaatcag	ttgttgatt	atccttatac	agtataggga	gatcacaata	3180
caactttatg	ccaataaaat	ctaacttaat	tgcccagata	tttttgcata	tttagcaaca	3240
agaaaagctt	atcatttgac	tcaagtttta	tgctttctct	ttcttttcat	ttcctaggta	3300
ctaattttta	tttttatttg	gaaggagcag	tgtaaagctt	acttgtattc	aatagtgtat	3360
ctcatagata	cagacaaggc	cgcagagata	agctgtttaa	tagtgtttaa	tgttgatgtg	3420
gagagaaagg	tgtattactt	aaaaatacta	taccatatac	gttttgata	tcattaaatc	3480
tttaaaagaa	attaaattta	ttcttgttta	aaaaaaaaaa	aaaaaaaaaa	aa	3532

<210> 8
 <211> 475
 <212> PRT
 <213> Mus Musculus

<400> 8

-7-

Ile	Met	Ala	Tyr	His	Glu	Lys	Pro	Val	Leu	Pro	Pro	Pro	Leu	Ile	Ile
1				5					10				15		
Leu	Ser	His	Ile	Val	Ser	Leu	Phe	Cys	Cys	Val	Cys	Lys	Arg	Arg	Lys
			20					25					30		
Lys	Asp	Lys	Thr	Ser	Asp	Gly	Pro	Lys	Leu	Phe	Leu	Thr	Glu	Glu	Asp
		35				40						45			
Gln	Lys	Lys	Leu	His	Asp	Phe	Glu	Glu	Gln	Cys	Val	Glu	Met	Tyr	Phe
	50					55					60				
Asp	Glu	Lys	Asp	Asp	Lys	Phe	Asn	Ser	Gly	Ser	Glu	Glu	Arg	Ile	Arg
65					70					75				80	
Val	Thr	Phe	Glu	Arg	Val	Glu	Gln	Met	Ser	Ile	Gln	Ile	Lys	Glu	Val
				85					90					95	
Gly	Asp	Arg	Val	Asn	Tyr	Ile	Lys	Arg	Ser	Leu	Gln	Ser	Leu	Asp	Ser
			100					105					110		
Gln	Ile	Gly	His	Leu	Gln	Asp	Leu	Ser	Ala	Leu	Thr	Val	Asp	Thr	Leu
		115					120					125			
Lys	Thr	Leu	Thr	Ala	Gln	Lys	Ala	Ser	Glu	Ala	Ser	Lys	Val	His	Asn
	130					135					140				
Glu	Ile	Thr	Arg	Glu	Leu	Ser	Ile	Ser	Lys	His	Leu	Ala	Gln	Asn	Leu
145					150					155					160
Ile	Asp	Asp	Val	Pro	Val	Arg	Pro	Leu	Trp	Glu	Glu	Pro	Ser	Ala	Val
				165					170					175	
Asn	Thr	Leu	Ser	Ser	Ser	Leu	Pro	Gln	Gly	Asp	Arg	Glu	Ser	Asn	Asn
			180					185					190		
Pro	Phe	Leu	Cys	Asn	Ile	Phe	Met	Lys	Asp	Glu	Lys	Asp	Pro	Gln	Tyr
		195					200					205			
Asn	Leu	Phe	Gly	Gln	Asp	Leu	Pro	Val	Ile	Pro	Gln	Arg	Lys	Glu	Phe
	210					215					220				
Asn	Ile	Pro	Glu	Ala	Gly	Ser	Ser	Cys	Gly	Ala	Leu	Phe	Pro	Ser	Ala
225					230					235					240
Val	Ser	Pro	Pro	Glu	Leu	Arg	Gln	Arg	Arg	His	Gly	Val	Glu	Met	Leu
				245					250					255	
Lys	Ile	Phe	Asn	Lys	Asn	Gln	Lys	Leu	Gly	Ser	Ser	Pro	Asn	Ser	Ser
			260					265					270		
Pro	His	Met	Ser	Ser	Pro	Pro	Thr	Lys	Phe	Ser	Val	Ser	Thr	Pro	Ser
		275					280					285			
Gln	Pro	Ser	Cys	Lys	Ser	His	Leu	Glu	Ser	Thr	Thr	Lys	Asp	Gln	Glu
	290					295					300				
Pro	Ile	Phe	Tyr	Lys	Ala	Ala	Glu	Gly	Asp	Asn	Ile	Glu	Phe	Gly	Ala
305					310					315					320
Phe	Val	Gly	His	Arg	Asp	Ser	Met	Asp	Leu	Gln	Arg	Phe	Lys	Glu	Thr
				325					330					335	
Ser	Asn	Lys	Ile	Arg	Glu	Leu	Leu	Ser	Asn	Asp	Thr	Pro	Glu	Asn	Thr
			340					345					350		
Leu	Lys	His	Val	Gly	Ala	Ala	Gly	Tyr	Ser	Glu	Cys	Cys	Lys	Thr	Ser
		355					360					365			
Thr	Ser	Leu	His	Ser	Val	Gln	Ala	Glu	Ser	Cys	Ser	Arg	Arg	Ala	Ser
						375					380				
Thr	Glu	Asp	Ser	Pro	Glu	Val	Asp	Ser	Lys	Ala	Ala	Leu	Leu	Pro	Asp
385					390					395					400
Trp	Leu	Arg	Asp	Arg	Pro	Ser	Asn	Arg	Glu	Met	Pro	Ser	Glu	Gly	Gly
				405					410					415	
Thr	Leu	Asn	Gly	Leu	Ala	Ser	Pro	Phe	Lys	Pro	Val	Leu	Asp	Thr	Asn
			420					425					430		
Tyr	Tyr	Tyr	Ser	Ala	Val	Glu	Arg	Asn	Asn	Leu	Met	Arg	Leu	Ser	Gln
		435					440					445			
Ser	Ile	Pro	Phe	Val	Pro	Val	Pro	Pro	Arg	Gly	Glu	Pro	Val	Thr	Val
		450				455					460				
Tyr	Pro	Ser	Gly	Gly	Arg	Val	Leu	Pro	Val	Tyr					
465					470					475					

<210> 9
 <211> 5433
 <212> DNA
 <213> Mus Musculus

<220>
 <221> unsure
 <222> (5094) ... (5094)
 <223> unknown

<400> 9
 ggctgaaaga gcctgagctg tgcctctcca ttccactgct gtggcagggg cagaaatctt 60
 ggatagagaa aaccttttgc aaacgggaat gtatctttgt aattcctagc acgaaagact 120
 ctaacagggtg ttgctgtggc cagttcacca accagcatat cccccctctg ccaagtgcga 180
 caccagcaaa aaatgaagag gaaagcaaac aggtggagac tcagcctgag aaatggctctg 240
 ttgccaaagca caccagagc taccacaacag attcctatgg agttcttgaa ttccaggggtg 300
 gcggatatcc caataaagcc atgtatatcc gtgtatccta tgacaccaag ccagactcac 360
 tgctccatct catggtgaaa gattggcagc tggaaactccc caagctctta atatctgtgc 420
 atggaggcct ccagaacttt gagatgcagc ccaagctgaa acaagtcctt gggaaaggcc 480
 tgatcaaggc tgctatgacc accggggcct ggatcttcac cgggggtgtc agcacagggtg 540
 ttatcagcca cgtaggggat gccttgaaa accactcctc caagtccaga ggccgggttt 600
 gtgctatagg aattgctcca tggggcatcg tggagaataa ggaagacctg gttggaaagg 660
 atgtaacaag agtgaccag accatgtcca accctctaag taagctctct gtgctcaaca 720
 actcccacac ccacttcac ctggctgaca atggcacctt gggcaagtat ggcgccgag 780
 tgaagctgctg aaggctgctg gaaaagcaca tctccctcca gaagatcaac acaagactgg 840
 ggcagggcgt gccctcgtg ggtctcgtg tggagggggg ccctaactgt gtgtccatcg 900
 tcttggaata cctgcaagaa gagcctccca tccctgtggt gatttgtgat ggcagcggac 960
 gtgctcggga catcctgtcc ttgcgacaca agtactgtga agaaggcggg ataataaatg 1020
 agtccctcag ggagcagctt ctagttaacca ttcagaaaaa atttaattat aataaggcac 1080
 aatcacatca gctgtttgca attataatgg agtgcattga gaagaaagaa ctctgactcg 1140
 tgttcagaat gggttctgag ggccagcagg acatcgagat ggcaatttta actgccctgc 1200
 tgaaaggaac aaacgtatct gctccagatc agctgagctt ggcactggct tggaaaccgg 1260
 tggacatagc acgaagccag atctttgtct ttgggcccc ctggacgccc ctgggaagcc 1320
 tggcaccccc gacggacagc aaagccacgg agaaggagaa gaagccacc atggccacca 1380
 ccaagggagg aagaggaaaa gggaaaggca agaagaaagg gaaagtgaag gaggaagtgg 1440
 aggaagaaac tgacccccgg aagatagagc tgctgaactg ggtgaatgct ttggagcaag 1500
 cgatgctaga tgctttagtc ttgactcgtg tcgactttgt gaagctcctg attgaaaaacg 1560
 gagtgaacat gcaacacttt ctgaccattc cgaggctgga ggagctctat aacacaagac 1620
 tgggtccacc aaacacactt catctgctgg tgagggatgt gaaaaagagc aaccttcgc 1680
 ctgattacca catcagcctc atagacatcg ggctcgtgct ggagtacctc atgggaggag 1740
 cctaccgctg caactacact cggaaaaact ttcggaccct ttacaacaac ttgtttggac 1800
 caaagaggcc taaagctctt aaacttctgg gaatggaaga tgatgagcct ccagctaaag 1860
 ggaagaaaaa aaaaaaaaag aaaaaggagg aagagatcga cattgatgtg gacgaccctg 1920
 ccgtgagtcg gttccagtat cccttccacg agctgatggt gtgggcagtg ctgatgaaac 1980
 gccagaaaat ggcagtgttc ctctggcagc gaggggaaga gagcatggcc aaggccctgg 2040
 tggcctgcaa gctctacaag gccatggccc acgagtcctc cgagagtgat ctggtggatg 2100
 acatctccca ggacttggat aacaattcca aagacttcgg ccagcttgct ttggagtta 2160
 tagaccagtc ctataagcat gacgagcaga tcgctatgaa actcctgacc tacgagctga 2220
 aaaactggag caactcgacc tgcctcaaac tggccgtggc agccaaacac cgggacttca 2280
 ttgctcacac ctgcagccag atgctgctga ccgatatgtg gatgggaaga ctgcggatgc 2340
 ggaagaaccc cggcctgaag gttatcatgg ggattcttct acccccacc atcttgtttt 2400
 tggaaatttcg cacatatgat gatttctcgt atcaaaccat caaggaaaac gaggatggca 2460
 aagaaaaaga agaggaaaat acggatgcaa atgcagatgc tggctcaaga aagggggatg 2520
 aggagaacga gcataaaaaa cagagaagta ttcccatcgg aacaaagatc tgtgaattct 2580
 ataacgcgcc cattgtcaag ttctggtttt acacaatatc atacttgggc tacctgtgce 2640
 tgtttaacta cgtcactcgt gtgcggatgg atggctggcc gtccctccag gattggatcg 2700
 tcatctccta catcgtgagc ctggcgttag agaagatacg agagatcctc atgtcagaac 2760
 caggcaaac cagccagaaa atcaaagttt ggcttcagga gtactggaac atcacagatc 2820
 tcgtggccat ttccacattc atgattggag caattcttcg cctacagaac cagccctaca 2880

-9-

```

tgggctatgg cggggtgatc tactgtgtgg atatcatctt ctggtacatc cgtgtcctgg 2940
acatctttgg tgtcaacaag tatctggggc catacgtgat gatgattgga aagatgatga 3000
tcgacatgct gtactttgtg gtcacatgc tggtcgtgct catgagtttc ggagtagccc 3060
gtcaagccat tctgcatcca gaggagaagc cctcttggaa actggcccga aacatcttct 3120
acatgcccta ctggatgatc tatggagagg tgtttgcaga ccagatagac ctctacgcca 3180
tggaatttaa tcctccttgt ggtgagaacc tatatgatga ggagggcaag cggcttcctc 3240
cctgtatccc cggcgccctg ctactccag cactcatggc gtgctatcta ctggtcgcca 3300
acatcctgct ggtgaacctg ctgattgctg tgttcaacaa tactttcttt gaagtaaaat 3360
caatatccaa ccaggtgtgg aagttccagc gatatcagct gattatgaca ttcatgaca 3420
ggccagtcct gccccaccg atgatcattt taagccacat ctacatcatc attatgcgtc 3480
tcagcggccg ctgcaggaaa aagagagaag gggaccaaga ggaacgggat cgtggattga 3540
agctcttctc tagcgacgag gagctaaaga ggctgcata gttcgaggag cagtgcgtgc 3600
aggagcactt ccgggagaag gaggatgagc agcagtcgtc cagcgacgag cgcatccggg 3660
tcacttctga aagagttgaa aatatgtcaa tgaggttgga agaaatcaat gaaagagaaa 3720
cttttatgaa aacttcctcg cagactgttg accttcgact tgctcagcta gaagaattat 3780
ctaacagaat ggtgaatgct cttgaaaatc ttgcgggaat cgacaggtct gacctgatcc 3840
aggcacggtc ccgggcttct tctgaatgtg aggcaacgta tcttctccgg caaagcagca 3900
tcaatagcgc tgatggctac agcttgatc gatatcattt taacggagaa gagttattat 3960
ttgaggatac atctctctcc acgtcaccag ggacaggagt caggaaaaaa acctgttcct 4020
tccgtataaa ggaagagaag gacgtgaaa cgcacctagt cccagaatgt cagaacagtc 4080
ttcaccttcc actgggcaca agcacatcag caaccctaga tggcagtcac cttgcagtag 4140
atgacttaaa gaacgctgaa gagtcaaaat taggtccaga tattgggatt tcaaaggaag 4200
atgatgaaag acagacagac tctaaaaaag aagaaactat ttccccaagt ttaaataaaa 4260
cagatgtgat acatggacag gacaaatcag atgttcaaaa cactcagcta acagtggaaa 4320
cgacaaatat agaaggcact atttcctatc ccctggaaga aacccaaatt acacgctatt 4380
tccccgatga aacgatcaat gcttgtaaaa caatgaagtc cagaagcttc gtctattccc 4440
ggggaagaaa gctggtcggt ggggttaacc aggatgtaga gtacagttca atcacggacc 4500
agcaattgac gacggaatgg caatgccaag ttcaaaaagat caccgctct catagcacag 4560
atattcctta cattgtgtcg gaagctgcag tgcaagctga gcaaaaagag cagtttgtag 4620
atatgcaaga tgaacacat gtcgctgaag caattcctcg aatccctcgc ttgtccctaa 4680
ccattactga cagaaatggg atggaaaact tactgtctgt gaagccagat caaactttgg 4740
gattcccatc tctcaggtca aaaagtttac atggacatcc taggaatgtg aaatccattc 4800
agggaaagt agacagatct ggacatgcca gtagtgaag cagcttagta attgtgtctg 4860
gaatgacagc agaagaaaaa aaggttaaga aagagaaaagc ttccacagaa actgaatgct 4920
agctgtgttt gtttctttaa ttttttttt taacagtcag aaaccacta atgggtgtca 4980
tcttgcccca tctaaacac atmtccaatt tcttaaaaac attttccctt aaaaaatttt 5040
ggaaattcag acttgattta caatttaatg cactaaaagt agtattttgt tagnatattg 5100
tagtaggctt agttttttca gttgcagtag tatcaaatga aagtgatgat actgtaacga 5160
agataaattg gctaatacag atacaagatt atacaatctc tttattactg agggccacca 5220
aatagcctag gaagtgcctc cgagcactga agtcaccatt aggtcactca agaagtaagc 5280
aactagctgg gcacagtggc tcatgcctgt aatcctagca ctttgggagg ccaaggcaga 5340
aagatagctt gagtccagga gtttgagacc agcctgggca acatagtgat acccatctc 5400
ttaaaaaaaa aaaaaaaaaa ctgcctcgt gcc 5433

```

<210> 10
 <211> 1533
 <212> PRT
 <213> Mus Musculus

<400> 10
 Met Tyr Ile Arg Val Ser Tyr Asp Thr Lys Pro Asp Ser Leu Leu His
 1 5 10 15
 Leu Met Val Lys Asp Trp Gln Leu Glu Leu Pro Lys Leu Leu Ile Ser
 20 25 30
 Val His Gly Gly Leu Gln Asn Phe Glu Met Gln Pro Lys Leu Lys Gln
 35 40 45
 Val Phe Gly Lys Gly Leu Ile Lys Ala Ala Met Thr Thr Gly Ala Trp
 50 55 60

-10-

Ile	Phe	Thr	Gly	Gly	Val	Ser	Thr	Gly	Val	Ile	Ser	His	Val	Gly	Asp	65	70	75	80
Ala	Leu	Lys	Asp	His	Ser	Ser	Lys	Ser	Arg	Gly	Arg	Val	Cys	Ala	Ile	85	90	95	
Gly	Ile	Ala	Pro	Trp	Gly	Ile	Val	Glu	Asn	Lys	Glu	Asp	Leu	Val	Gly	100	105	110	
Lys	Asp	Val	Thr	Arg	Val	Tyr	Gln	Thr	Met	Ser	Asn	Pro	Leu	Ser	Lys	115	120	125	
Leu	Ser	Val	Leu	Asn	Asn	Ser	His	Thr	His	Phe	Ile	Leu	Ala	Asp	Asn	130	135	140	
Gly	Thr	Leu	Gly	Lys	Tyr	Gly	Ala	Glu	Val	Lys	Leu	Arg	Arg	Leu	Leu	145	150	155	160
Glu	Lys	His	Ile	Ser	Leu	Gln	Lys	Ile	Asn	Thr	Arg	Leu	Gly	Gln	Gly	165	170	175	
Val	Pro	Leu	Val	Gly	Leu	Val	Val	Glu	Gly	Gly	Pro	Asn	Val	Val	Ser	180	185	190	
Ile	Val	Leu	Glu	Tyr	Leu	Gln	Glu	Glu	Pro	Pro	Ile	Pro	Val	Val	Ile	195	200	205	
Cys	Asp	Gly	Ser	Gly	Arg	Ala	Ser	Asp	Ile	Leu	Ser	Phe	Ala	His	Lys	210	215	220	
Tyr	Cys	Glu	Glu	Gly	Gly	Ile	Ile	Asn	Glu	Ser	Leu	Arg	Glu	Gln	Leu	225	230	235	240
Leu	Val	Thr	Ile	Gln	Lys	Thr	Phe	Asn	Tyr	Asn	Lys	Ala	Gln	Ser	His	245	250	255	
Gln	Leu	Phe	Ala	Ile	Ile	Met	Glu	Cys	Met	Lys	Lys	Lys	Glu	Leu	Val	260	265	270	
Thr	Val	Phe	Arg	Met	Gly	Ser	Glu	Gly	Gln	Gln	Asp	Ile	Glu	Met	Ala	275	280	285	
Ile	Leu	Thr	Ala	Leu	Leu	Lys	Gly	Thr	Asn	Val	Ser	Ala	Pro	Asp	Gln	290	295	300	
Leu	Ser	Leu	Ala	Leu	Ala	Trp	Asn	Arg	Val	Asp	Ile	Ala	Arg	Ser	Gln	305	310	315	320
Ile	Phe	Val	Phe	Gly	Pro	His	Trp	Thr	Pro	Leu	Gly	Ser	Leu	Ala	Pro	325	330	335	
Pro	Thr	Asp	Ser	Lys	Ala	Thr	Glu	Lys	Glu	Lys	Lys	Pro	Pro	Met	Ala	340	345	350	
Thr	Thr	Lys	Gly	Gly	Arg	Gly	Lys	Gly	Lys	Gly	Lys	Lys	Lys	Gly	Lys	355	360	365	
Val	Lys	Glu	Glu	Val	Glu	Glu	Glu	Thr	Asp	Pro	Arg	Lys	Ile	Glu	Leu	370	375	380	
Leu	Asn	Trp	Val	Asn	Ala	Leu	Glu	Gln	Ala	Met	Leu	Asp	Ala	Leu	Val	385	390	395	400
Leu	Asp	Arg	Val	Asp	Phe	Val	Lys	Leu	Leu	Ile	Glu	Asn	Gly	Val	Asn	405	410	415	
Met	Gln	His	Phe	Leu	Thr	Ile	Pro	Arg	Leu	Glu	Glu	Leu	Tyr	Asn	Thr	420	425	430	
Arg	Leu	Gly	Pro	Pro	Asn	Thr	Leu	His	Leu	Leu	Val	Arg	Asp	Val	Lys	435	440	445	
Lys	Ser	Asn	Leu	Pro	Pro	Asp	Tyr	His	Ile	Ser	Leu	Ile	Asp	Ile	Gly	450	455	460	
Leu	Val	Leu	Glu	Tyr	Leu	Met	Gly	Gly	Ala	Tyr	Arg	Cys	Asn	Tyr	Thr	465	470	475	480
Arg	Lys	Asn	Phe	Arg	Thr	Leu	Tyr	Asn	Asn	Leu	Phe	Gly	Pro	Lys	Arg	485	490	495	
Pro	Lys	Ala	Leu	Lys	Leu	Leu	Gly	Met	Glu	Asp	Asp	Glu	Pro	Pro	Ala	500	505	510	
Lys	Gly	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Glu	Glu	Glu	Ile	Asp	Ile	515	520	525	
Asp	Val	Asp	Asp	Pro	Ala	Val	Ser	Arg	Phe	Gln	Tyr	Pro	Phe	His	Glu	530	535	540	

-11-

Leu	Met	Val	Trp	Ala	Val	Leu	Met	Lys	Arg	Gln	Lys	Met	Ala	Val	Phe
545					550					555					560
Leu	Trp	Gln	Arg	Gly	Glu	Glu	Ser	Met	Ala	Lys	Ala	Leu	Val	Ala	Cys
				565					570						575
Lys	Leu	Tyr	Lys	Ala	Met	Ala	His	Glu	Ser	Ser	Glu	Ser	Asp	Leu	Val
			580					585					590		
Asp	Asp	Ile	Ser	Gln	Asp	Leu	Asp	Asn	Asn	Ser	Lys	Asp	Phe	Gly	Gln
		595					600					605			
Leu	Ala	Leu	Glu	Leu	Leu	Asp	Gln	Ser	Tyr	Lys	His	Asp	Glu	Gln	Ile
	610					615					620				
Ala	Met	Lys	Leu	Leu	Thr	Tyr	Glu	Leu	Lys	Asn	Trp	Ser	Asn	Ser	Thr
625					630					635					640
Cys	Leu	Lys	Leu	Ala	Val	Ala	Ala	Lys	His	Arg	Asp	Phe	Ile	Ala	His
				645					650					655	
Thr	Cys	Ser	Gln	Met	Leu	Leu	Thr	Asp	Met	Trp	Met	Gly	Arg	Leu	Arg
			660					665					670		
Met	Arg	Lys	Asn	Pro	Gly	Leu	Lys	Val	Ile	Met	Gly	Ile	Leu	Leu	Pro
		675					680					685			
Pro	Thr	Ile	Leu	Phe	Leu	Glu	Phe	Arg	Thr	Tyr	Asp	Asp	Phe	Ser	Tyr
	690					695					700				
Gln	Thr	Ser	Lys	Glu	Asn	Glu	Asp	Gly	Lys	Glu	Lys	Glu	Glu	Glu	Asn
705					710					715					720
Thr	Asp	Ala	Asn	Ala	Asp	Ala	Gly	Ser	Arg	Lys	Gly	Asp	Glu	Glu	Asn
				725					730					735	
Glu	His	Lys	Lys	Gln	Arg	Ser	Ile	Pro	Ile	Gly	Thr	Lys	Ile	Cys	Glu
			740					745					750		
Phe	Tyr	Asn	Ala	Pro	Ile	Val	Lys	Phe	Trp	Phe	Tyr	Thr	Ile	Ser	Tyr
		755					760					765			
Leu	Gly	Tyr	Leu	Leu	Leu	Phe	Asn	Tyr	Val	Ile	Leu	Val	Arg	Met	Asp
	770						775					780			
Gly	Trp	Pro	Ser	Leu	Gln	Glu	Trp	Ile	Val	Ile	Ser	Tyr	Ile	Val	Ser
785					790					795					800
Leu	Ala	Leu	Glu	Lys	Ile	Arg	Glu	Ile	Leu	Met	Ser	Glu	Pro	Gly	Lys
				805					810					815	
Leu	Ser	Gln	Lys	Ile	Lys	Val	Trp	Leu	Gln	Glu	Tyr	Trp	Asn	Ile	Thr
		820						825					830		
Asp	Leu	Val	Ala	Ile	Ser	Thr	Phe	Met	Ile	Gly	Ala	Ile	Leu	Arg	Leu
		835					840					845			
Gln	Asn	Gln	Pro	Tyr	Met	Gly	Tyr	Gly	Arg	Val	Ile	Tyr	Cys	Val	Asp
	850					855					860				
Ile	Ile	Phe	Trp	Tyr	Ile	Arg	Val	Leu	Asp	Ile	Phe	Gly	Val	Asn	Lys
865					870					875					880
Tyr	Leu	Gly	Pro	Tyr	Val	Met	Met	Ile	Gly	Lys	Met	Met	Ile	Asp	Met
				885					890					895	
Leu	Tyr	Phe	Val	Val	Ile	Met	Leu	Val	Val	Leu	Met	Ser	Phe	Gly	Val
		900						905					910		
Ala	Arg	Gln	Ala	Ile	Leu	His	Pro	Glu	Glu	Lys	Pro	Ser	Trp	Lys	Leu
		915					920					925			
Ala	Arg	Asn	Ile	Phe	Tyr	Met	Pro	Tyr	Trp	Met	Ile	Tyr	Gly	Glu	Val
		930				935					940				
Phe	Ala	Asp	Gln	Ile	Asp	Leu	Tyr	Ala	Met	Glu	Ile	Asn	Pro	Pro	Cys
945					950					955					960
Gly	Glu	Asn	Leu	Tyr	Asp	Glu	Glu	Gly	Lys	Arg	Leu	Pro	Pro	Cys	Ile
			965						970					975	
Pro	Gly	Ala	Trp	Leu	Thr	Pro	Ala	Leu	Met	Ala	Cys	Tyr	Leu	Leu	Val
			980					985					990		
Ala	Asn	Ile	Leu	Leu	Val	Asn	Leu	Leu	Ile	Ala	Val	Phe	Asn	Asn	Thr
		995					1000					1005			
Phe	Phe	Glu	Val	Lys	Ser	Ile	Ser	Asn	Gln	Val	Trp	Lys	Phe	Gln	Arg
1010						1015					1020				

-12-

Tyr Gln Leu Ile Met Thr Phe His Asp Arg Pro Val Leu Pro Pro Pro
 1025 1030 1035 104
 Met Ile Ile Leu Ser His Ile Tyr Ile Ile Ile Met Arg Leu Ser Gly
 1045 1050 1055
 Arg Cys Arg Lys Lys Arg Glu Gly Asp Gln Glu Glu Arg Asp Arg Gly
 1060 1065 1070
 Leu Lys Leu Phe Leu Ser Asp Glu Glu Leu Lys Arg Leu His Glu Phe
 1075 1080 1085
 Glu Glu Gln Cys Val Gln Glu His Phe Arg Glu Lys Glu Asp Glu Gln
 1090 1095 1100
 Gln Ser Ser Ser Asp Glu Arg Ile Arg Val Thr Ser Glu Arg Val Glu
 1105 1110 1115 112
 Asn Met Ser Met Arg Leu Glu Glu Ile Asn Glu Arg Glu Thr Phe Met
 1125 1130 1135
 Lys Thr Ser Leu Gln Thr Val Asp Leu Arg Leu Ala Gln Leu Glu Glu
 1140 1145 1150
 Leu Ser Asn Arg Met Val Asn Ala Leu Glu Asn Leu Ala Gly Ile Asp
 1155 1160 1165
 Arg Ser Asp Leu Ile Gln Ala Arg Ser Arg Ala Ser Ser Glu Cys Glu
 1170 1175 1180
 Ala Thr Tyr Leu Leu Arg Gln Ser Ser Ile Asn Ser Ala Asp Gly Tyr
 1185 1190 1195 120
 Ser Leu Tyr Arg Tyr His Phe Asn Gly Glu Glu Leu Leu Phe Glu Asp
 1205 1210 1215
 Thr Ser Leu Ser Thr Ser Pro Gly Thr Gly Val Arg Lys Lys Thr Cys
 1220 1225 1230
 Ser Phe Arg Ile Lys Glu Glu Lys Asp Val Lys Thr His Leu Val Pro
 1235 1240 1245
 Glu Cys Gln Asn Ser Leu His Leu Ser Leu Gly Thr Ser Thr Ser Ala
 1250 1255 1260
 Thr Pro Asp Gly Ser His Leu Ala Val Asp Asp Leu Lys Asn Ala Glu
 1265 1270 1275 128
 Glu Ser Lys Leu Gly Pro Asp Ile Gly Ile Ser Lys Glu Asp Asp Glu
 1285 1290 1295
 Arg Gln Thr Asp Ser Lys Lys Glu Glu Thr Ile Ser Pro Ser Leu Asn
 1300 1305 1310
 Lys Thr Asp Val Ile His Gly Gln Asp Lys Ser Asp Val Gln Asn Thr
 1315 1320 1325
 Gln Leu Thr Val Glu Thr Thr Asn Ile Glu Gly Thr Ile Ser Tyr Pro
 1330 1335 1340
 Leu Glu Glu Thr Lys Ile Thr Arg Tyr Phe Pro Asp Glu Thr Ile Asn
 1345 1350 1355 136
 Ala Cys Lys Thr Met Lys Ser Arg Ser Phe Val Tyr Ser Arg Gly Arg
 1365 1370 1375
 Lys Leu Val Gly Gly Val Asn Gln Asp Val Glu Tyr Ser Ser Ile Thr
 1380 1385 1390
 Asp Gln Gln Leu Thr Thr Glu Trp Gln Cys Gln Val Gln Lys Ile Thr
 1395 1400 1405
 Arg Ser His Ser Thr Asp Ile Pro Tyr Ile Val Ser Glu Ala Ala Val
 1410 1415 1420
 Gln Ala Glu Gln Lys Glu Gln Phe Ala Asp Met Gln Asp Glu His His
 1425 1430 1435 144
 Val Ala Glu Ala Ile Pro Arg Ile Pro Arg Leu Ser Leu Thr Ile Thr
 1445 1450 1455
 Asp Arg Asn Gly Met Glu Asn Leu Leu Ser Val Lys Pro Asp Gln Thr
 1460 1465 1470
 Leu Gly Phe Pro Ser Leu Arg Ser Lys Ser Leu His Gly His Pro Arg
 1475 1480 1485
 Asn Val Lys Ser Ile Gln Gly Lys Leu Asp Arg Ser Gly His Ala Ser
 1490 1495 1500

-13-

Ser Val Ser Ser Leu Val Ile Val Ser Gly Met Thr Ala Glu Glu Lys
 1505 1510 1515 152
 Lys Val Lys Lys Glu Lys Ala Ser Thr Glu Thr Glu Cys
 1525 1530

<210> 11
 <211> 6220
 <212> DNA
 <213> Homo Sapiens

<400> 11

tgtgcagaat	tgtacagttg	cgaaacccatg	tcgctggcag	ctgggtgctgg	cgggtggagac	60
ttccctgtgc	ggtgctcagt	gcatctgcac	ccgtggggga	gggagctctt	tctctggccc	120
tgcagtcacc	tgaggttggt	accattatga	acggccgctg	ggacccccgc	atgtgcatgt	180
actccccag	agtgtccggg	ggccccagcc	aagggaacaca	tctcacgcag	ctgggaacat	240
gtgcaggctg	atgaagagaa	ccggatgagg	gcttcacatg	aggaagcatg	tggccaggtc	300
ctctcagaac	atcagccctca	tcttcctgtc	tctgatctat	ttcaccaacc	accccatgtg	360
tctctagaac	cccagtgtag	cgagctggag	agaggactgt	cctgagggca	gcaggcctgg	420
ttgcagctgg	cgtgggggtc	tcagaatgga	gccctcagcc	ctgaggaaag	ctggctcgga	480
gcaggaggag	ggctttgagg	ggctgcccag	aagggtcact	gacctgggga	tggcttccaa	540
tctccggcgc	agcaacagca	gcctcttcaa	gagctggagg	ctacagtggc	ccttcggcaa	600
caatgacaag	caagaaagcc	tcagttcgtg	gattcctgaa	aacatcaaga	agaaagaatg	660
cgtgtatttt	gtggaaagtt	ccaaactgtc	tgatgctggg	aagggtggtg	gtcagtgtgg	720
ctacacgcat	gagcagcact	tggaggaggc	taccaagccc	cacaccttcc	agggcacaca	780
gtgggaccca	aagaaacatg	tccaggagat	gccaaaccgat	gcctttggcg	acatcgtctt	840
cacgggcctg	agccagaagg	tgaaaaagta	cgctccgagt	tcccaggaca	cgccctccag	900
cgtgatctac	cacctcatga	cccagcactg	ggggctggac	gtccccaatc	tcttgatctc	960
ggtgaccggg	ggggccaaga	acttcaacat	gaagccgcgg	ctgaagagca	ttttccgcag	1020
aggcctggtc	aagggtggctc	agaccacagg	ggcctggatc	atcacagggg	ggtcccacac	1080
cggcgtcatg	aagcaggtag	gcgaggcggg	gcgggacttc	agcctgagca	gcagctacaa	1140
ggaaggcgag	ctcatcacca	tcggagtcgc	caacctggggc	actgtccacc	gccgcgaggg	1200
cctgatccat	cccacgggca	gcttccccgc	cgagtacata	ctggatgagg	atggccaagg	1260
gaacctgacc	tgcctagaca	gcaaccactc	tacttccatc	ctcgtggacg	acgggaccca	1320
cggccagtag	ggggtggaga	ttcctctgag	gaccaggctg	gagaagttca	tatcggagca	1380
gaccaaggaa	agaggagggtg	tggccatcaa	gatccccatc	gtgtgctgtg	tgctggaggg	1440
cggcccgggc	acgttgacaca	ccatcgacaa	cgccaaccac	aacggcacc	cctgtgtggt	1500
tgtggagggc	tcgggccgcg	tggccgacgt	cattgccag	gtggccaacc	tgcctgtctc	1560
ggacatcact	atctccctga	tccagcagaa	actgagcgtg	ttcttccagg	agatgtttga	1620
gaccttcacg	gaaagcagga	ttgtcagagt	gacaaaaaag	atccaagata	ttgtccggag	1680
gcggcagctg	ctgactgtct	tccgggaagg	caaggatggg	cagcaggacg	tggatgtggc	1740
catcttgcag	gccttgctga	aagcctcacg	gagccaagac	cactttggcc	acgagaactg	1800
ggaccaccag	ctgaaactgg	cagtggcatg	gaatcgctg	gacattgccc	gcagtgcagt	1860
cttcatggat	gagtggcagt	ggaagccttc	agatctgcac	cccacgatga	cagctgcact	1920
catctccaac	aagcctgagt	ttgtgaagct	cttctgggaa	aacgggggtg	agctgaaggga	1980
gtttgtcacc	tgggacacct	tgctctacct	gtacgagaac	ctggacccct	cctgcctgtt	2040
ccacagcaag	ctgcaaaagg	tgctggtgga	ggatcccag	cgcccggctt	gcgcgcccg	2100
ggcggccccg	ctgcagatgc	accacgtggc	ccagggtgct	cgggagctgc	tgggggactt	2160
cacgcagccg	ctttatcccc	ggccccggca	caacgaccgg	ctgcggctcc	tgctgcccgt	2220
tccccacgtc	aagctcaacg	tgcagggagt	gagcctccgg	tccctctaca	agcgttcctc	2280
aggccatgtg	accttcacca	tggaccccat	ccgtgacctt	ctcatttggg	ccattgtcca	2340
gaaccgtcgg	gagctggcag	gaatcatctg	ggctcagagc	caggactgca	tcgcagcgcc	2400
cttggcctgc	agcaagatcc	tgaaggaaact	gtccaaggag	gaggaggaca	cggacagctc	2460
ggaggagatg	ctggcgctgg	cggaggagta	tgagcacaga	gccatcgggg	tcttcaccga	2520
gtgctaccgg	aaggacgaag	agagagccca	gaaactgtct	acccgcgtgt	ccgaggcctg	2580
ggggaagacc	acctgcctgc	agctcgccct	ggaggccaag	gacatgaagt	ttgtgtctca	2640
cgggggcatc	caggccttcc	tgaccaaggt	gtgggtgggg	cagctctccg	tggacaattg	2700
gctgtggcgt	gtgacctgtg	gcatgctggc	cttcccgtg	ctcctcaccg	gcctcatctc	2760
cttcaggagg	aagaggtgct	aggaatgtgg	caccccgcg	gcccgcgccc	gtgccttctt	2820
caccgcacc	gtggtggtct	tccacctgaa	cactctctcc	tacttcgcct	tcctctgcct	2880
gttcgcctac	gtgctcatgg	tggacttcca	gcctgtgccc	tcctggtgcg	agtgtgccat	2940

ctacctctgg	ctcttctcct	tgggtgtgca	ggagatgcgg	cagctcttct	atgaccctga	3000
cgagtgcggg	ctgatgaaga	aggcagcctt	gracttcagt	gacttctgga	ataagctgga	3060
cgtcggcgca	atcttgctct	tcgtggcagg	gctgacctgc	aggctcatcc	cggcgacgct	3120
gtaccccggg	cgcgctcatcc	tctctctgga	cttcatcctg	ttctgcctcc	ggctcatgca	3180
catttttacc	atcagtaaga	cgctggggcc	caagatcatc	attgtgaagc	ggatgatgaa	3240
ggacgtcttc	ttcttctcct	tctgtctggc	tgtgtgggtg	gtgtccttcg	gggtggccaa	3300
gcaggccatc	ctcatccaca	acgagcgccg	ggtggactgg	ctgttccgag	gggcccgtcta	3360
ccactcctac	ctcaccatct	tcgggcagat	cccgggctac	atcgacggtg	tgaacttcaa	3420
cccggagcac	tgcagcccca	atggcaccga	cccctacaag	cctaagtgcc	ccgagagcga	3480
cgcgacgcag	cagaggcccg	ccttccctga	gtggctgacg	gtcctcctac	tctgcctcta	3540
cctgctcttc	accaacatcc	tgtgtctcaa	cctcctcatc	gccatgttca	actacacctt	3600
ccagcaggtg	caggagcaca	cggaccagat	ttggaagttc	cagcgccatg	acctgatcga	3660
ggagtaccac	ggccgccccg	ccgcgcgcgc	ccccttcctc	ctcctcagcc	acctgcagct	3720
cttcatcaag	aggggtggtcc	tgaagactcc	ggccaagagg	cacaagcagc	tcaagaacaa	3780
gctggagaag	aacgaggagg	cgccctgctt	atcctgggag	atctacctga	aggagaacta	3840
cctccagaag	cgacagttcc	agcaaaagca	gcggcccagc	cagaagatcg	aggacatcag	3900
caataagggt	gacggcatgg	tggacctgct	cgacctggac	ccactgaaga	ggtcgggctc	3960
catggagcag	aggttggcct	ccctggagga	gcaggtggcc	cagacagccc	gagccctgca	4020
ctggatcgtg	aggacgctgc	gggccagcgg	cttcagctcg	gaggcggacg	tccccactct	4080
ggcctcccag	aaggccgcgg	aggagccgga	tgttgagccg	ggaggcagga	agaagacgga	4140
ggagccgggc	gacagctacc	acgtgaatgc	ccggcacctc	ctctacccca	actgccctgt	4200
cacgcgcttc	cccgtgccca	acgagaaggt	gccctgggag	acggagtctc	tgatctatga	4260
cccacccttt	tacacggcag	agaggaagga	cgcgcccgcc	atggacccca	tgggagacac	4320
cctggagcca	ctgtccacga	tccagtacaa	cgtgggtgat	ggcctgaggg	accgccggag	4380
cttccacggg	ccgtacacag	tgcaggccgg	gttgccctcg	aaccccatgg	gccgcacagg	4440
actgcgtggg	cgcgggagcc	tcagctgctt	cggaccacaac	cacacgctgt	accccatggt	4500
cacgcggttg	agggcggaacg	aggatggagc	catctgcagg	aagagcataa	agaagatgct	4560
ggaagtgtcg	gtgggtgaagc	tccctctctc	cgagcactgg	gccctgcctg	ggggctcccg	4620
ggagccaggg	gagatgctac	ctcggaagct	gaagcggatc	ctccggcagg	agcactggcc	4680
gtcttttgaa	aacttgctga	agtgcggcat	ggagggtgac	aaaggctaca	tggatgaccc	4740
gaggaaacag	gacaatgcct	ggatcgagac	ggtggccgtc	agcgtccact	tccaggacca	4800
gaatgacgtg	gagctgaaca	ggctgaactc	taacctgcac	gcctgcgact	cgggggcctc	4860
catccgatgg	caggtggtgg	acaggcgcat	cccactctat	gcgaaccaca	agaccctcct	4920
ccagaaggca	gccgtgagt	tcggggctca	ctactgactg	tgccctcagg	ctgggcggtc	4980
ccagtccata	gacgttcccc	ccagaaacca	gggcttctct	ctcctgagcc	tggccaggac	5040
tcaggctggt	cctgggccct	gcacatgatg	gggtttgggt	gacccagtgc	ccctcacggc	5100
tgccgcaagt	ctgctgcaga	tgacctcatg	aactggaagg	ggtcaagggt	acccggggag	5160
agagctcaag	acagggcaca	ggctactcag	agctgagggg	cccctgggac	ccttggccat	5220
caggcgaggg	gctgggcctg	tgacgtggg	cccttggccca	gagtccactc	ccttccctgg	5280
tgtgtcacc	cgagcagctc	atccaccatg	gaggtcattg	gcctgaggca	agttccccgg	5340
agagtccgga	tcccctgtgg	ccccctcagg	cctatgtctg	tgaggaaggg	gccctgccac	5400
tctccccaag	agggcctcca	tgtttcgagg	tgccctcaaca	tggagccttg	cctggccttg	5460
gctaggggca	ctgtctgaac	tcctgactgt	caggataaac	tccgtggggg	tacaggagcc	5520
cagacaaagc	ccaggcctgt	caagagacgc	agaggggccc	tgccagggtt	ggccccaggg	5580
accctgggac	gaggctgcag	aagctctccc	tccctactcc	ctgggagcca	cgtgctggcc	5640
atgtggccag	ggacggcatg	agcaggaggc	ggggacgtgg	gggccttctg	gtttggtgtc	5700
aacagctcac	aggagcgtga	accatgaggg	ccctcaggag	gggaacgtgg	taaaacccaa	5760
gacattaaat	ctgccatctc	aggcctggct	ggctcttctg	tgctttccac	aaataaagtt	5820
cctgacacgt	ccaggggccag	gggctgtgtg	acggctgcct	gaagttctcc	tcgatcccc	5880
ggtgagcttc	ctgcagcctg	tggatgtcct	gcagccccct	agccctaccc	ccaagtttct	5940
cctctgaccc	atcagctccc	tgtcttcatt	ttcctaaacc	tgggctccag	catcgtcccc	6000
aagccacca	ggccaggatg	caggcatcca	catgccctcc	tccttggtct	cccctgcgtg	6060
gtggtgccaa	tgtgccctg	cacccctgca	gaggtccgg	atggagcctg	gggctgcctg	6120
gccactgagc	actggccgag	gtgatgcccc	cccttccctg	gacaggcctc	tgtcttccac	6180
ctgacccaaa	gctctctagc	caccccttgc	tccccagtat			6220

<210> 12

<211> 1503

<212> PRT

<213> Homo Sapiens

<400> 12

Met	Glu	Pro	Ser	Ala	Leu	Arg	Lys	Ala	Gly	Ser	Glu	Gln	Glu	Glu	Gly
1				5					10					15	
Phe	Glu	Gly	Leu	Pro	Arg	Arg	Val	Thr	Asp	Leu	Gly	Met	Val	Ser	Asn
			20					25					30		
Leu	Arg	Arg	Ser	Asn	Ser	Ser	Leu	Phe	Lys	Ser	Trp	Arg	Leu	Gln	Cys
			35				40					45			
Pro	Phe	Gly	Asn	Asn	Asp	Lys	Gln	Glu	Ser	Leu	Ser	Ser	Trp	Ile	Pro
			50			55					60				
Glu	Asn	Ile	Lys	Lys	Lys	Glu	Cys	Val	Tyr	Phe	Val	Glu	Ser	Ser	Lys
					70					75					80
Leu	Ser	Asp	Ala	Gly	Lys	Val	Val	Cys	Gln	Cys	Gly	Tyr	Thr	His	Glu
				85					90					95	
Gln	His	Leu	Glu	Glu	Ala	Thr	Lys	Pro	His	Thr	Phe	Gln	Gly	Thr	Gln
			100					105					110		
Trp	Asp	Pro	Lys	Lys	His	Val	Gln	Glu	Met	Pro	Thr	Asp	Ala	Phe	Gly
			115				120					125			
Asp	Ile	Val	Phe	Thr	Gly	Leu	Ser	Gln	Lys	Val	Lys	Lys	Tyr	Val	Arg
			130			135					140				
Val	Ser	Gln	Asp	Thr	Pro	Ser	Ser	Val	Ile	Tyr	His	Leu	Met	Thr	Gln
					150					155					160
His	Trp	Gly	Leu	Asp	Val	Pro	Asn	Leu	Leu	Ile	Ser	Val	Thr	Gly	Gly
				165				170						175	
Ala	Lys	Asn	Phe	Asn	Met	Lys	Pro	Arg	Leu	Lys	Ser	Ile	Phe	Arg	Arg
			180					185					190		
Gly	Leu	Val	Lys	Val	Ala	Gln	Thr	Thr	Gly	Ala	Trp	Ile	Ile	Thr	Gly
			195				200					205			
Gly	Ser	His	Thr	Gly	Val	Met	Lys	Gln	Val	Gly	Glu	Ala	Val	Arg	Asp
			210			215					220				
Phe	Ser	Leu	Ser	Ser	Ser	Tyr	Lys	Glu	Gly	Glu	Leu	Ile	Thr	Ile	Gly
					230					235					240
Val	Ala	Thr	Trp	Gly	Thr	Val	His	Arg	Arg	Glu	Gly	Leu	Ile	His	Pro
				245				250						255	
Thr	Gly	Ser	Phe	Pro	Ala	Glu	Tyr	Ile	Leu	Asp	Glu	Asp	Gly	Gln	Gly
			260					265					270		
Asn	Leu	Thr	Cys	Leu	Asp	Ser	Asn	His	Ser	His	Phe	Ile	Leu	Val	Asp
			275				280					285			
Asp	Gly	Thr	His	Gly	Gln	Tyr	Gly	Val	Glu	Ile	Pro	Leu	Arg	Thr	Arg
			290			295					300				
Leu	Glu	Lys	Phe	Ile	Ser	Glu	Gln	Thr	Lys	Glu	Arg	Gly	Gly	Val	Ala
					310					315					320
Ile	Lys	Ile	Pro	Ile	Val	Cys	Val	Val	Leu	Glu	Gly	Gly	Pro	Gly	Thr
				325					330					335	
Leu	His	Thr	Ile	Asp	Asn	Ala	Thr	Thr	Asn	Gly	Thr	Pro	Cys	Val	Val
			340					345					350		
Val	Glu	Gly	Ser	Gly	Arg	Val	Ala	Asp	Val	Ile	Ala	Gln	Val	Ala	Asn
			355				360					365			
Leu	Pro	Val	Ser	Asp	Ile	Thr	Ile	Ser	Leu	Ile	Gln	Gln	Lys	Leu	Ser
			370			375					380				
Val	Phe	Phe	Gln	Glu	Met	Phe	Glu	Thr	Phe	Thr	Glu	Ser	Arg	Ile	Val
					390					395					400
Glu	Trp	Thr	Lys	Lys	Ile	Gln	Asp	Ile	Val	Arg	Arg	Arg	Gln	Leu	Leu
				405					410					415	
Thr	Val	Phe	Arg	Glu	Gly	Lys	Asp	Gly	Gln	Gln	Asp	Val	Asp	Val	Ala
			420					425					430		
Ile	Leu	Gln	Ala	Leu	Leu	Lys	Ala	Ser	Arg	Ser	Gln	Asp	His	Phe	Gly
			435				440					445			

-16-

His	Glu	Asn	Trp	Asp	His	Gln	Leu	Lys	Leu	Ala	Val	Ala	Trp	Asn	Arg
450						455					460				
Val	Asp	Ile	Ala	Arg	Ser	Glu	Ile	Phe	Met	Asp	Glu	Trp	Gln	Trp	Lys
465					470					475					480
Pro	Ser	Asp	Leu	His	Pro	Thr	Met	Thr	Ala	Ala	Leu	Ile	Ser	Asn	Lys
				485					490					495	
Pro	Glu	Phe	Val	Lys	Leu	Phe	Leu	Glu	Asn	Gly	Val	Gln	Leu	Lys	Glu
			500					505					510		
Phe	Val	Thr	Trp	Asp	Thr	Leu	Leu	Tyr	Leu	Tyr	Glu	Asn	Leu	Asp	Pro
		515				520					525				
Ser	Cys	Leu	Phe	His	Ser	Lys	Leu	Gln	Lys	Val	Leu	Val	Glu	Asp	Pro
	530					535					540				
Glu	Arg	Pro	Ala	Cys	Ala	Pro	Ala	Ala	Pro	Arg	Leu	Gln	Met	His	His
545					550					555					560
Val	Ala	Gln	Val	Leu	Arg	Glu	Leu	Leu	Gly	Asp	Phe	Thr	Gln	Pro	Leu
				565					570					575	
Tyr	Pro	Arg	Pro	Arg	His	Asn	Asp	Arg	Leu	Arg	Leu	Leu	Leu	Pro	Val
			580					585					590		
Pro	His	Val	Lys	Leu	Asn	Val	Gln	Gly	Val	Ser	Leu	Arg	Ser	Leu	Tyr
		595					600					605			
Lys	Arg	Ser	Ser	Gly	His	Val	Thr	Phe	Thr	Met	Asp	Pro	Ile	Arg	Asp
	610					615					620				
Leu	Leu	Ile	Trp	Ala	Ile	Val	Gln	Asn	Arg	Arg	Glu	Leu	Ala	Gly	Ile
625					630					635					640
Ile	Trp	Ala	Gln	Ser	Gln	Asp	Cys	Ile	Ala	Ala	Ala	Leu	Ala	Cys	Ser
				645					650					655	
Lys	Ile	Leu	Lys	Glu	Leu	Ser	Lys	Glu	Glu	Glu	Asp	Thr	Asp	Ser	Ser
			660					665					670		
Glu	Glu	Met	Leu	Ala	Leu	Ala	Glu	Glu	Tyr	Glu	His	Arg	Ala	Ile	Gly
		675					680					685			
Val	Phe	Thr	Glu	Cys	Tyr	Arg	Lys	Asp	Glu	Glu	Arg	Ala	Gln	Lys	Leu
	690					695					700				
Leu	Thr	Arg	Val	Ser	Glu	Ala	Trp	Gly	Lys	Thr	Cys	Leu	Gln	Leu	
705					710					715					720
Ala	Leu	Glu	Ala	Lys	Asp	Met	Lys	Phe	Val	Ser	His	Gly	Gly	Ile	Gln
				725					730					735	
Ala	Phe	Leu	Thr	Lys	Val	Trp	Trp	Gly	Gln	Leu	Ser	Val	Asp	Asn	Gly
			740					745					750		
Leu	Trp	Arg	Val	Thr	Leu	Cys	Met	Leu	Ala	Phe	Pro	Leu	Leu	Leu	Thr
		755					760					765			
Gly	Leu	Ile	Ser	Phe	Arg	Glu	Lys	Arg	Leu	Gln	Asp	Val	Gly	Thr	Pro
	770					775					780				
Ala	Ala	Arg	Ala	Arg	Ala	Phe	Phe	Thr	Ala	Pro	Val	Val	Val	Phe	His
785					790					795					800
Leu	Asn	Ile	Leu	Ser	Tyr	Phe	Ala	Phe	Leu	Cys	Leu	Phe	Ala	Tyr	Val
				805					810					815	
Leu	Met	Val	Asp	Phe	Gln	Pro	Val	Pro	Ser	Trp	Cys	Glu	Cys	Ala	Ile
			820					825					830		
Tyr	Leu	Trp	Leu	Phe	Ser	Leu	Val	Cys	Glu	Glu	Met	Arg	Gln	Leu	Phe
		835					840					845			
Tyr	Asp	Pro	Asp	Glu	Cys	Gly	Leu	Met	Lys	Lys	Ala	Ala	Leu	Tyr	Phe
	850					855					860				
Ser	Asp	Phe	Trp	Asn	Lys	Leu	Asp	Val	Gly	Ala	Ile	Leu	Leu	Phe	Val
865					870					875					880
Ala	Gly	Leu	Thr	Cys	Arg	Leu	Ile	Pro	Ala	Thr	Leu	Tyr	Pro	Gly	Arg
				885					890					895	
Val	Ile	Leu	Ser	Leu	Asp	Phe	Ile	Leu	Phe	Cys	Leu	Arg	Leu	Met	His
		900						905					910		
Ile	Phe	Thr	Ile	Ser	Lys	Thr	Leu	Gly	Pro	Lys	Ile	Ile	Ile	Val	Lys
		915					920						925		

-17-

Arg	Met	Met	Lys	Asp	Val	Phe	Phe	Phe	Leu	Phe	Leu	Leu	Ala	Val	Trp
930						935					940				
Val	Val	Ser	Phe	Gly	Val	Ala	Lys	Gln	Ala	Ile	Leu	Ile	His	Asn	Glu
945					950					955					960
Arg	Arg	Val	Asp	Trp	Leu	Phe	Arg	Gly	Ala	Val	Tyr	His	Ser	Tyr	Leu
			965						970					975	
Thr	Ile	Phe	Gly	Gln	Ile	Pro	Gly	Tyr	Ile	Asp	Gly	Val	Asn	Phe	Asn
		980						985					990		
Pro	Glu	His	Cys	Ser	Pro	Asn	Gly	Thr	Asp	Pro	Tyr	Lys	Pro	Lys	Cys
		995					1000					1005			
Pro	Glu	Ser	Asp	Ala	Thr	Gln	Gln	Arg	Pro	Ala	Phe	Pro	Glu	Trp	Leu
	1010					1015					1020				
Thr	Val	Leu	Leu	Leu	Cys	Leu	Tyr	Leu	Leu	Phe	Thr	Asn	Ile	Leu	Leu
1025					1030					1035					104
Leu	Asn	Leu	Leu	Ile	Ala	Met	Phe	Asn	Tyr	Thr	Phe	Gln	Gln	Val	Gln
				1045					1050					1055	
Glu	His	Thr	Asp	Gln	Ile	Trp	Lys	Phe	Gln	Arg	His	Asp	Leu	Ile	Glu
		1060						1065					1070		
Glu	Tyr	His	Gly	Arg	Pro	Ala	Ala	Pro	Pro	Pro	Phe	Ile	Leu	Leu	Ser
	1075					1080					1085				
His	Leu	Gln	Leu	Phe	Ile	Lys	Arg	Val	Val	Leu	Lys	Thr	Pro	Ala	Lys
	1090					1095					1100				
Arg	His	Lys	Gln	Leu	Lys	Asn	Lys	Leu	Glu	Lys	Asn	Glu	Glu	Ala	Ala
1105					1110					1115					112
Leu	Leu	Ser	Trp	Glu	Ile	Tyr	Leu	Lys	Glu	Asn	Tyr	Leu	Gln	Asn	Arg
			1125						1130					1135	
Gln	Phe	Gln	Gln	Lys	Gln	Arg	Pro	Glu	Gln	Lys	Ile	Glu	Asp	Ile	Ser
		1140						1145					1150		
Asn	Lys	Val	Asp	Ala	Met	Val	Asp	Leu	Leu	Asp	Leu	Asp	Pro	Leu	Lys
	1155						1160				1165				
Arg	Ser	Gly	Ser	Met	Glu	Gln	Arg	Leu	Ala	Ser	Leu	Glu	Glu	Gln	Val
	1170					1175					1180				
Ala	Gln	Thr	Ala	Arg	Ala	Leu	His	Trp	Ile	Val	Arg	Thr	Leu	Arg	Ala
1185					1190					1195					120
Ser	Gly	Phe	Ser	Ser	Glu	Ala	Asp	Val	Pro	Thr	Leu	Ala	Ser	Gln	Lys
			1205						1210					1215	
Ala	Ala	Glu	Glu	Pro	Asp	Ala	Glu	Pro	Gly	Gly	Arg	Lys	Lys	Thr	Glu
	1220						1225						1230		
Glu	Pro	Gly	Asp	Ser	Tyr	His	Val	Asn	Ala	Arg	His	Leu	Leu	Tyr	Pro
	1235						1240					1245			
Asn	Cys	Pro	Val	Thr	Arg	Phe	Pro	Val	Pro	Asn	Glu	Lys	Val	Pro	Trp
	1250					1255					1260				
Glu	Thr	Glu	Phe	Leu	Ile	Tyr	Asp	Pro	Pro	Phe	Tyr	Thr	Ala	Glu	Arg
1265					1270					1275					128
Lys	Asp	Ala	Ala	Ala	Met	Asp	Pro	Met	Gly	Asp	Thr	Leu	Glu	Pro	Leu
			1285						1290					1295	
Ser	Thr	Ile	Gln	Tyr	Asn	Val	Val	Asp	Gly	Leu	Arg	Asp	Arg	Arg	Ser
		1300						1305					1310		
Phe	His	Gly	Pro	Tyr	Thr	Val	Gln	Ala	Gly	Leu	Pro	Leu	Asn	Pro	Met
	1315						1320					1325			
Gly	Arg	Thr	Gly	Leu	Arg	Gly	Arg	Gly	Ser	Leu	Ser	Cys	Phe	Gly	Pro
	1330					1335					1340				
Asn	His	Thr	Leu	Tyr	Pro	Met	Val	Thr	Arg	Trp	Arg	Arg	Asn	Glu	Asp
1345					1350					1355					136
Gly	Ala	Ile	Cys	Arg	Lys	Ser	Ile	Lys	Lys	Met	Leu	Glu	Val	Leu	Val
			1365						1370					1375	
Val	Lys	Leu	Pro	Leu	Ser	Glu	His	Trp	Ala	Leu	Pro	Gly	Gly	Ser	Arg
		1380						1385					1390		
Glu	Pro	Gly	Glu	Met	Leu	Pro	Arg	Lys	Leu	Lys	Arg	Ile	Leu	Arg	Gln
	1395						1400					1405			

-18-

Glu His Trp Pro Ser Phe Glu Asn Leu Leu Lys Cys Gly Met Glu Val
 1410 1415 1420
 Tyr Lys Gly Tyr Met Asp Asp Pro Arg Asn Thr Asp Asn Ala Trp Ile
 1425 1430 1435 144
 Glu Thr Val Ala Val Ser Val His Phe Gln Asp Gln Asn Asp Val Glu
 1445 1450 1455
 Leu Asn Arg Leu Asn Ser Asn Leu His Ala Cys Asp Ser Gly Ala Ser
 1460 1465 1470
 Ile Arg Trp Gln Val Val Asp Arg Arg Ile Pro Leu Tyr Ala Asn His
 1475 1480 1485
 Lys Thr Leu Leu Gln Lys Ala Ala Ala Glu Phe Gly Ala His Tyr
 1490 1495 1500

<210> 13
 <211> 1816
 <212> PRT
 <213> C. Elegans

<400> 13
 Met Ile Thr Asp Lys Asn Leu Phe Ser Arg Leu Leu Ile Lys Lys Asn
 1 5 10 15
 Pro Ile Arg Met His Ser Pro Ser Phe Ser Phe Ser Leu Ile Thr Ser
 20 25 30
 Leu Phe Phe Thr Gln Phe Phe Met Phe Gln Leu Ser Ser Met Ala Tyr
 35 40 45
 Phe Phe Leu Thr Leu Ile Ala Gly Val Thr His Phe Tyr Phe Pro Glu
 50 55 60
 Lys Leu Leu Gly Lys Ser Glu Asn Leu Asp His Arg Tyr Gln Ser Ser
 65 70 75 80
 Glu Gln Lys Val Leu Ile Glu Trp Thr Glu Asn Lys Ala Val Ala Glu
 85 90 95
 Ser Leu Arg Ala Asn Ser Val Thr Val Glu Glu Asn Glu Ser Glu Arg
 100 105 110
 Glu Thr Glu Thr Gln Thr Lys Arg Arg Arg Lys Lys Gln Arg Ser Thr
 115 120 125
 Ser Ser Asp Lys Ala Pro Leu Asn Ser Ala Pro Arg His Val Gln Lys
 130 135 140
 Phe Asp Trp Lys Asp Met Leu His Leu Ala Asp Ile Ser Gly Arg Lys
 145 150 155 160
 Arg Gly Asn Ser Thr Thr Ser His Ser Gly His Ala Thr Arg Ala Gly
 165 170 175
 Ser Leu Lys Gly Lys Asn Trp Ile Glu Cys Arg Leu Lys Met Arg Gln
 180 185 190
 Cys Ser Tyr Phe Val Pro Ser Gln Arg Phe Ser Glu Arg Cys Gly Cys
 195 200 205
 Gly Lys Glu Arg Ser Lys His Thr Glu Glu Val Leu Glu Arg Ser Gln
 210 215 220
 Asn Lys Asn His Pro Leu Asn His Leu Thr Leu Pro Gly Ile His Glu
 225 230 235 240
 Val Asp Thr Thr Asp Ala Asp Ala Asp Asp Asn Glu Val Asn Leu Thr
 245 250 255
 Pro Gly Arg Trp Ser Ile Gln Ser His Thr Glu Ile Val Pro Thr Asp
 260 265 270
 Ala Tyr Gly Asn Ile Val Phe Glu Gly Thr Ala His His Ala Gln Tyr
 275 280 285
 Ala Arg Ile Ser Phe Asp Ser Asp Pro Arg Asp Ile Val His Leu Met
 290 295 300
 Met Lys Val Trp Lys Leu Lys Pro Pro Lys Leu Ile Ile Thr Ile Asn
 305 310 315 320
 Gly Gly Leu Thr Lys Phe Asp Leu Gln Pro Lys Leu Ala Arg Thr Phe

				325					330					335	
Arg	Lys	Gly	Ile	Met	Lys	Ile	Ala	Lys	Ser	Thr	Asp	Ala	Trp	Ile	Ile
			340					345					350		
Thr	Ser	Gly	Leu	Asp	Glu	Gly	Val	Val	Lys	His	Leu	Asp	Ser	Ala	Leu
		355					360					365			
His	Ala	Leu	Glu	Phe	Trp	Ser	Phe	Gly	Leu	Phe	Trp	Val	Ile	Gln	Leu
	370					375				380					
Asp	Val	Leu	Leu	Ala	His	Ser	Met	Phe	Ile	Pro	Arg	Gly	Ser	Leu	Phe
385					390					395					400
Asp	His	Gly	Asn	His	Thr	Ser	Lys	Asn	His	Val	Val	Ala	Ile	Gly	Ile
			405						410					415	
Ala	Ser	Trp	Gly	Met	Leu	Lys	Gln	Arg	Ser	Arg	Phe	Val	Gly	Lys	Asp
			420					425					430		
Ser	Thr	Val	Thr	Tyr	Ala	Thr	Asn	Val	Phe	Asn	Asn	Thr	Arg	Leu	Lys
	435					440					445				
Glu	Leu	Asn	Asp	Asn	His	Ser	Tyr	Phe	Leu	Phe	Ser	Asp	Asn	Gly	Thr
	450					455					460				
Val	Asn	Arg	Tyr	Gly	Ala	Glu	Ile	Ile	Met	Arg	Lys	Arg	Leu	Glu	Ala
465					470					475					480
Tyr	Leu	Ala	Gln	Gly	Asp	Lys	Lys	Arg	Ser	Ala	Ile	Pro	Leu	Val	Cys
			485						490					495	
Val	Val	Leu	Glu	Gly	Gly	Ala	Phe	Thr	Ile	Lys	Met	Val	His	Asp	Tyr
			500					505					510		
Val	Thr	Thr	Ile	Pro	Arg	Ile	Pro	Val	Ile	Val	Cys	Asp	Gly	Ser	Gly
	515						520					525			
Arg	Ala	Ala	Asp	Ile	Leu	Ala	Phe	Ala	His	Gln	Ala	Val	Ser	Gln	Asn
	530					535					540				
Gly	Phe	Leu	Ser	Asp	Asn	Ile	Arg	Asn	Gln	Leu	Val	Asn	Ile	Val	Arg
545					550					555					560
Arg	Ile	Phe	Gly	Tyr	Asp	Pro	Lys	Thr	Ala	Gln	Lys	Leu	Ile	Lys	Gln
			565						570					575	
Ile	Val	Glu	Cys	Ser	Thr	Asn	Lys	Ser	Leu	Met	Thr	Ile	Phe	Arg	Leu
			580					585					590		
Gly	Glu	Ser	Ser	Arg	Glu	Asp	Leu	Asp	His	Val	Ile	Met	Ser	Cys	Leu
	595						600					605			
Leu	Lys	Gly	Gln	Asn	Leu	Ser	Pro	Pro	Glu	Gln	Leu	Gln	Leu	Ala	Leu
	610					615					620				
Ala	Trp	Asn	Arg	Ala	Asp	Ile	Ala	Arg	Thr	Glu	Ile	Phe	Ala	Asn	Gly
625					630					635					640
Thr	Glu	Trp	Thr	Thr	Gln	Asp	Leu	His	Asn	Ala	Met	Ile	Glu	Ala	Leu
			645						650					655	
Ser	Asn	Asp	Arg	Ile	Asp	Phe	Val	His	Leu	Leu	Leu	Glu	Asn	Gly	Val
			660					665					670		
Ser	Met	Gln	Lys	Phe	Leu	Thr	Tyr	Gly	Arg	Leu	Glu	His	Leu	Tyr	Asn

					805					810					815
Asn	Met	Asp	Phe	Thr	Phe	Arg	Tyr	Pro	Tyr	Ser	Asp	Leu	Met	Ile	Trp
			820					825					830		
Ala	Val	Leu	Thr	Lys	Arg	Gln	Lys	Met	Ala	Lys	Leu	Met	Trp	Thr	His
		835					840					845			
Gly	Glu	Glu	Gly	Met	Ala	Lys	Ala	Leu	Val	Ala	Ser	Arg	Leu	Tyr	Val
	850					855					860				
Ser	Leu	Ala	Lys	Thr	Ala	Ser	Leu	Ala	Thr	Gly	Glu	Ile	Gly	Met	Ser
865				870						875				880	
Gln	Asp	Phe	Thr	Glu	Phe	Ser	Asp	Glu	Phe	Ser	Glu	Leu	Ala	Val	Glu
			885						890				895		
Val	Leu	Glu	Tyr	Cys	Thr	Lys	His	Gly	Arg	Asp	Gln	Thr	Leu	Arg	Leu
		900						905					910		
Leu	Thr	Cys	Glu	Leu	Ala	Asn	Trp	Gly	Asp	Glu	Thr	Cys	Leu	Ser	Leu
		915					920					925			
Ala	Ala	Asn	Asn	Gly	His	Arg	Lys	Phe	Leu	Ala	His	Pro	Cys	Cys	Gln
	930					935					940				
Met	Leu	Leu	Ser	Asp	Leu	Trp	Gln	Gly	Gly	Leu	Leu	Met	Lys	Asn	Asn
945				950					955					960	
Gln	Asn	Ser	Lys	Val	Leu	Thr	Cys	Leu	Ala	Ala	Pro	Pro	Leu	Ile	Phe
			965						970				975		
Leu	Leu	Gly	Phe	Lys	Thr	Lys	Glu	Gln	Leu	Met	Leu	Gln	Pro	Lys	Thr
		980					985					990			
Ala	Ala	Glu	His	Asp	Glu	Glu	Met	Ser	Asp	Ser	Glu	Met	Asn	Ser	Ala
		995				1000						1005			
Glu	Asp	Thr	Asp	Thr	Ser	Ser	Asp	Ser	Ser	Ser	Asp	Ser	Asp	Asp	Ser
	1010					1015					1020				
Asp	Glu	Glu	Asp	Ala	Lys	Leu	Arg	Ala	Gln	Ser	Leu	Ser	Ala	Asp	Gln
1025			1030						1035					104	
Pro	Leu	Ser	Ile	His	Arg	Leu	Val	Arg	Asp	Lys	Leu	Asn	Phe	Ser	Glu
			1045						1050					1055	
Lys	Lys	Lys	Pro	Asp	Met	Gly	Ile	Ser	Arg	Ile	Val	Val	Ala	Pro	Pro
		1060					1065					1070			
Ile	Val	Thr	Gly	Arg	Asn	Arg	Ala	Arg	Thr	Met	Ser	Ile	Lys	Lys	Ser
	1075					1080						1085			
Lys	Lys	Asn	Val	Ile	Lys	Pro	Pro	Ala	Cys	Leu	Lys	Ile	Glu	Thr	Ser
	1090					1095					1100				
Asp	Asp	Asp	Glu	Gln	Glu	Gln	Lys	Lys	Ala	Thr	Glu	Met	Cys	Lys	Ser
1105			1110						1115					112	
Thr	Phe	Phe	Asp	Phe	Phe	Phe	Asp	Phe	Pro	Tyr	Ile	Asn	Arg	Thr	Gly
			1125						1130					1135	
Lys	Arg	Gly	Ser	Val	Ala	Val	Ala	Met	Asn	His	Asp	Asp	Met	Tyr	Ile
		1140					1145					1150			
Asp	Pro	Ser	Glu	Glu	Leu	Asp	Thr	Gln	Thr	Arg	Gln	Lys			

-21-

				1285					1290				1295		
Thr	Thr	Lys	Thr	Leu	Gly	Arg	Ile	Leu	Ile	Ile	Cys	Asn	Ser	Val	Ile
				1300					1305					1310	
Trp	Ser	Leu	Lys	Leu	Val	Asp	Tyr	Leu	Ser	Val	Gln	Gln	Gly	Leu	Gly
				1315					1320					1325	
Pro	Tyr	Ile	Asn	Ile	Val	Ala	Glu	Met	Ile	Pro	Thr	Met	Ile	Pro	Leu
				1330					1335					1340	
Cys	Val	Leu	Val	Phe	Ile	Thr	Leu	Tyr	Ala	Phe	Gly	Leu	Leu	Arg	Gln
1345							1350				1355				136
Ser	Ile	Thr	Tyr	Pro	Tyr	Glu	Asp	Trp	His	Trp	Ile	Leu	Val	Arg	Asn
				1365					1370						1375
Ile	Phe	Leu	Gln	Pro	Tyr	Phe	Met	Leu	Tyr	Gly	Glu	Val	Tyr	Ala	Ala
				1380					1385						1390
Glu	Ile	Asp	Thr	Cys	Gly	Asp	Glu	Ile	Trp	Gln	Thr	His	Glu	Asp	Glu
				1395					1400					1405	
Asn	Ile	Pro	Ile	Ser	Met	Leu	Asn	Val	Thr	His	Glu	Thr	Cys	Val	Pro
				1410					1415						1420
Gly	Tyr	Trp	Ile	Ala	Pro	Val	Gly	Leu	Thr	Val	Phe	Met	Leu	Ala	Thr
1425							1430				1435				144
Asn	Val	Leu	Leu	Met	Asn	Val	Met	Val	Ala	Gly	Cys	Thr	Tyr	Ile	Phe
				1445					1450						1455
Glu	Lys	His	Ile	Gln	Ser	Thr	Arg	Glu	Ile	Phe	Leu	Phe	Glu	Arg	Tyr
				1460					1465					1470	
Gly	Gln	Val	Met	Glu	Tyr	Glu	Ser	Thr	Pro	Trp	Leu	Pro	Pro	Pro	Phe
				1475					1480					1485	
Thr	Ile	Ile	Tyr	His	Val	Ile	Trp	Leu	Phe	Lys	Leu	Ile	Lys	Ser	Ser
				1490					1495					1500	
Ser	Arg	Met	Phe	Glu	Arg	Lys	Asn	Leu	Phe	Asp	Gln	Ser	Leu	Lys	Leu
1505							1510				1515				152
Phe	Leu	Ser	Pro	Asp	Glu	Met	Glu	Lys	Val	His	Thr	Phe	Glu	Glu	Glu
				1525					1530						1535
Ser	Val	Glu	Asp	Met	Lys	Arg	Glu	Thr	Glu	Lys	Lys	Asn	Leu	Ser	Ser
				1540					1545						1550
Asn	Asp	Glu	Arg	Ile	His	Arg	Thr	Ala	Glu	Arg	Thr	Asp	Ala	Ile	Leu
				1555					1560						1565
Asn	Arg	Val	Ser	His	Leu	Thr	Gln	Leu	Glu	Phe	Thr	Leu	Lys	Glu	Glu
				1570					1575						1580
Ile	Arg	Glu	Leu	Glu	His	Lys	Met	Lys	Asn	Met	Asp	Ser	Arg	His	Lys
1585							1590				1595				160
Glu	Gln	Met	Asn	Leu	Met	Leu	Asp	Met	Asn	Lys	Lys	Leu	Gly	Lys	Phe
				1605					1610						1615
Ile	Ser	Gly	Lys	Tyr	Lys	Arg	Gly	Ser	Phe	Gly	Gly	Ser	Gly	Ser	Asp
				1620					1625						1630
Gly	Gly	Gly	Gly	Ser	Ser	Asp	Asn	Ser	Lys	Leu	Glu	Pro	Asn	Asn	Ser
				1635					1640						1645
Val	Pro	Met	Ile	Thr	Val	Asp	Gly	Pro	Ser	Pro	Ile	Gly	Ser	Arg	Arg
				1650					1655						1660
Thr	Ser	Gly	Gln	Tyr	Leu	Lys	Arg	Asp	Ser	Leu	Gln	Ala	Lys	Lys	Lys
1665							1670				1675				168
Ile	Thr	Glu	Asn	Arg	Arg	Ser	Ser	Leu	Glu	Gln	Pro	Lys	Ile	Pro	Ser
				1685					1690						1695
Ile	Gln	Phe	Asn	Leu	Met	Glu	Asp	Gln	Asp	Glu	Ser	Ala	Ala	Glu	Ser
				1700					1705						1710
Ala	Thr	Glu	Glu	Val	Ser	Ile	Ser	Ile	Pro	Val	Pro	Gln	Met	Arg	Val
				1715					1720						1725
Arg	Gln	Val	Thr	Glu	Ser	Asp	Lys	Ser	Asp	Leu	Ser	Glu	Asp	Asp	Leu
				1730					1735						1740
Ile	Thr	Arg	Glu	Asp	Ala	Pro	Pro	Thr	Ser	Ile	Asn	Leu	Pro	Arg	Gly
1745							1750				1755				176
Pro	Arg	Arg	His	Ala	Leu	Tyr	Ser	Thr	Ile	Ala	Asp	Ala	Ile	Glu	Thr

-22-

			1765						1770					1775
Glu	Asp	Asp	Phe	Tyr	Ala	Asp	Ser	Pro	Val	Pro	Met	Pro	Met	Thr
			1780						1785				1790	
Val	Gln	Pro	Ala	Asp	Gly	Ser	Phe	Phe	Gly	Glu	Asn	Asp	Ser	Arg
			1795				1800					1805		
Gln	Arg	Asp	Asp	Ser	Asp	Tyr	Glu							
	1810					1815								

<210> 14
 <211> 1387
 <212> PRT
 <213> C. Elegans

<400> 14

Met	Arg	Lys	Ser	Arg	Arg	Val	Arg	Lys	Leu	Val	Arg	His	Ala	Ser	Leu
1			5						10					15	
Ile	Glu	Asn	Ile	Arg	His	Arg	Thr	Ser	Ser	Phe	Leu	Arg	Leu	Leu	Asn
			20					25					30		
Ala	Pro	Arg	Asn	Ser	Met	Cys	Asn	Ala	Asn	Thr	Val	His	Ser	Ile	Ser
		35				40						45			
Ser	Phe	Arg	Ser	Asp	His	Leu	Ser	Arg	Lys	Ser	Thr	His	Lys	Phe	Leu
	50				55						60				
Asp	Asn	Pro	Asn	Leu	Phe	Ala	Ile	Glu	Leu	Thr	Glu	Lys	Leu	Ser	Pro
	65			70				75						80	
Pro	Trp	Ile	Glu	Asn	Thr	Phe	Glu	Lys	Arg	Glu	Cys	Ile	Arg	Phe	Ala
			85					90						95	
Ala	Leu	Pro	Lys	Asp	Pro	Glu	Arg	Cys	Gly	Cys	Gly	Arg	Pro	Leu	Ser
			100					105					110		
Ala	His	Thr	Pro	Ala	Ser	Thr	Phe	Ser	Thr	Leu	Pro	Val	His	Leu	
	115					120						125			
Leu	Glu	Lys	Glu	Gln	Gln	Thr	Trp	Thr	Ile	Ala	Asn	Asn	Thr	Gln	Thr
	130					135					140				
Ser	Thr	Thr	Asp	Ala	Phe	Gly	Thr	Ile	Val	Phe	Gln	Gly	Gly	Ala	His
	145			150						155				160	
Ala	His	Lys	Ala	Gln	Tyr	Val	Arg	Leu	Ser	Tyr	Asp	Ser	Glu	Pro	Leu
			165						170					175	
Asp	Val	Met	Tyr	Leu	Met	Glu	Lys	Val	Trp	Gly	Leu	Glu	Ala	Pro	Arg
		180						185					190		
Leu	Val	Ile	Thr	Val	His	Gly	Gly	Met	Ser	Asn	Phe	Glu	Leu	Glu	Glu
	195					200						205			
Arg	Leu	Gly	Arg	Leu	Phe	Arg	Lys	Gly	Met	Leu	Lys	Ala	Ala	Gln	Thr
	210					215					220				
Thr	Gly	Ala	Trp	Ile	Ile	Thr	Ser	Gly	Leu	Asp	Ser	Gly	Val	Val	Arg
	225			230						235				240	
His	Val	Ala	Lys	Ala	Leu	Asp	Glu	Ala	Gly	Ile	Ser	Ala	Arg	Met	Arg
			245						250					255	
Ser	Gln	Ile	Val	Thr	Ile	Gly	Ile	Ala	Pro	Trp	Gly	Val	Ile	Lys	Arg
			260					265					270		
Lys	Glu	Arg	Leu	Ile	Arg	Gln	Asn	Glu	His	Val	Tyr	Tyr	Asp	Val	His
	275					280						285			
Ser	Leu	Ser	Val	Asn	Ala	Asn	Val	Gly	Ile	Leu	Asn	Asp	Arg	His	Ser
	290					295					300				
Tyr	Phe	Leu	Leu	Ala	Asp	Asn	Gly	Thr	Val	Gly	Arg	Phe	Gly	Ala	Asp
	305			310						315				320	
Leu	His	Leu	Arg	Gln	Asn	Leu	Glu	Asn	His	Ile	Ala	Thr	Phe	Gly	Cys
			325						330					335	
Asn	Gly	Arg	Lys	Val	Pro	Val	Val	Cys	Thr	Leu	Leu	Glu	Gly	Gly	Ile
			340					345					350		
Ser	Ser	Ile	Asn	Ala	Ile	His	Asp	Tyr	Val	Thr	Met	Lys	Pro	Asp	Ile
		355				360						365			

-23-

Pro	Ala	Ile	Val	Cys	Asp	Gly	Ser	Gly	Arg	Ala	Ala	Asp	Ile	Ile	Ser
370						375				380					
Phe	Ala	Ala	Arg	Tyr	Ile	Asn	Ser	Asp	Gly	Thr	Phe	Ala	Ala	Glu	Val
385					390					395					400
Gly	Glu	Lys	Leu	Arg	Asn	Leu	Ile	Lys	Met	Val	Phe	Pro	Glu	Thr	Asp
				405					410						415
Gln	Glu	Glu	Met	Phe	Arg	Lys	Ile	Thr	Glu	Cys	Val	Ile	Arg	Asp	Asp
			420					425						430	
Leu	Leu	Arg	Ile	Phe	Arg	Tyr	Gly	Gln	Glu	Glu	Glu	Glu	Asp	Val	Asp
		435					440					445			
Phe	Val	Ile	Leu	Ser	Thr	Val	Leu	Gln	Lys	Gln	Asn	Leu	Pro	Pro	Asp
	450					455					460				
Glu	Gln	Leu	Ala	Leu	Thr	Leu	Ser	Trp	Asn	Arg	Val	Asp	Leu	Ala	Lys
465					470					475					480
Ser	Cys	Leu	Phe	Ser	Asn	Gly	Arg	Lys	Trp	Ser	Ser	Asp	Val	Leu	Glu
				485					490						495
Lys	Ala	Met	Asn	Asp	Ala	Leu	Tyr	Trp	Asp	Arg	Val	Asp	Phe	Val	Glu
			500					505					510		
Cys	Leu	Leu	Glu	Asn	Gly	Val	Ser	Met	Lys	Asn	Phe	Leu	Ser	Ile	Asn
		515					520					525			
Arg	Leu	Glu	Asn	Leu	Tyr	Asn	Met	Asp	Asp	Ile	Asn	Ser	Ala	His	Ser
		530				535						540			
Val	Arg	Asn	Trp	Met	Glu	Asn	Phe	Asp	Ser	Met	Asp	Pro	His	Thr	Tyr
545					550					555					560
Leu	Thr	Ile	Pro	Met	Ile	Gly	Gln	Val	Val	Glu	Lys	Leu	Met	Gly	Asn
				565					570						575
Ala	Phe	Gln	Leu	Tyr	Thr	Ser	Arg	Ser	Phe	Lys	Gly	Lys	Tyr	Asp	
			580				585					590			
Arg	Tyr	Lys	Arg	Ile	Asn	Gln	Ser	Ser	Tyr	Phe	His	Arg	Lys	Arg	Lys
		595					600					605			
Ile	Val	Gln	Lys	Glu	Leu	Phe	Lys	Lys	Lys	Ser	Asp	Asp	Gln	Ile	Asn
	610					615					620				
Asp	Asn	Glu	Glu	Glu	Asp	Phe	Ser	Phe	Ala	Tyr	Pro	Phe	Asn	Asp	Leu
625					630					635					640
Leu	Ile	Trp	Ala	Val	Leu	Thr	Ser	Arg	His	Gly	Met	Ala	Glu	Cys	Met
				645					650					655	
Trp	Val	His	Gly	Glu	Asp	Ala	Met	Ala	Lys	Cys	Leu	Leu	Ala	Ile	Arg
			660					665					670		
Leu	Tyr	Lys	Ala	Thr	Ala	Lys	Ile	Ala	Glu	Asp	Glu	Tyr	Leu	Asp	Val
		675					680					685			
Glu	Glu	Ala	Lys	Arg	Leu	Phe	Asp	Asn	Ala	Val	Lys	Cys	Arg	Glu	Asp
		690				695					700				
Ala	Ile	Glu	Leu	Leu	Asp	Gln	Cys	Tyr	Arg	Ala	Asp	His	Asp	Arg	Thr
705					710					715					720
Leu	Arg	Leu	Leu	Arg	Met	Glu	Leu	Pro	His	Trp	Gly	Asn	Asn	Asn	Cys
				725					730					735	
Leu	Ser	Leu	Ala	Val	Leu	Ala	Asn	Thr	Lys	Thr	Phe	Leu	Ala	His	Pro
			740					745					750		
Cys	Cys	Gln	Ile	Leu	Leu	Ala	Glu	Leu	Trp	His	Gly	Ser	Leu	Lys	Val
		755					760					765			
Arg	Ser	Gly	Ser	Asn	Val	Arg	Val	Leu	Thr	Ala	Leu	Ile	Cys	Pro	Pro
	770					775						780			
Ala	Ile	Leu	Phe	Met	Ala	Tyr	Lys	Pro	Lys	His	Ser	Lys	Thr	Ala	Arg
785					790					795					800
Leu	Leu	Ser	Glu	Glu	Thr	Pro	Glu	Gln	Leu	Pro	Tyr	Pro	Arg	Glu	Ser
				805					810					815	
Ile	Thr	Ser	Thr	Thr	Ser	Asn	Arg	Tyr	Arg	Tyr	Ser	Lys	Gly	Pro	Glu
			820					825					830		
Glu	Gln	Lys	Glu	Thr	Leu	Leu	Glu	Lys	Gly	Ser	Tyr	Thr	Lys	Lys	Val
		835					840					845			

-24-

Thr	Ile	Ile	Ser	Ser	Arg	Lys	Asn	Ser	Gly	Val	Ala	Ser	Val	Tyr	Gly	850	855	860
Ser	Ala	Ser	Ser	Met	Met	Phe	Lys	Arg	Glu	Pro	Gln	Leu	Asn	Lys	Phe	865	870	875
Glu	Arg	Phe	Arg	Ala	Phe	Tyr	Ser	Ser	Pro	Ile	Thr	Lys	Phe	Trp	Ser	885	890	895
Trp	Cys	Ile	Ala	Phe	Leu	Ile	Phe	Leu	Thr	Thr	Gln	Thr	Cys	Ile	Leu	900	905	910
Leu	Leu	Glu	Thr	Ser	Leu	Lys	Pro	Ser	Lys	Tyr	Glu	Trp	Ile	Thr	Phe	915	920	925
Ile	Tyr	Thr	Val	Thr	Leu	Ser	Val	Glu	His	Ile	Arg	Lys	Leu	Met	Thr	930	935	940
Ser	Glu	Gly	Ser	Arg	Ile	Asn	Glu	Lys	Val	Lys	Val	Phe	Tyr	Ala	Lys	945	950	955
Trp	Tyr	Asn	Ile	Trp	Thr	Ser	Ala	Ala	Leu	Leu	Phe	Phe	Leu	Val	Gly	965	970	975
Tyr	Gly	Phe	Arg	Leu	Val	Pro	Met	Tyr	Arg	His	Ser	Trp	Gly	Arg	Val	980	985	990
Leu	Leu	Ser	Phe	Ser	Asn	Val	Leu	Phe	Tyr	Met	Lys	Ile	Phe	Glu	Tyr	995	1000	1005
Leu	Ser	Val	His	Pro	Leu	Leu	Gly	Pro	Tyr	Ile	Gln	Met	Ala	Ala	Lys	1010	1015	1020
Met	Val	Trp	Ser	Met	Cys	Tyr	Ile	Cys	Val	Leu	Leu	Leu	Val	Pro	Leu	1025	1030	1035
Met	Ala	Phe	Gly	Val	Asn	Arg	Gln	Ala	Leu	Thr	Glu	Pro	Asn	Val	Lys	1045	1050	1055
Asp	Trp	His	Trp	Leu	Leu	Val	Arg	Asn	Ile	Phe	Tyr	Lys	Pro	Tyr	Phe	1060	1065	1070
Met	Leu	Tyr	Gly	Glu	Val	Tyr	Ala	Gly	Glu	Ile	Asp	Thr	Cys	Gly	Asp	1075	1080	1085
Glu	Gly	Ile	Arg	Cys	Phe	Pro	Gly	Tyr	Phe	Ile	Pro	Pro	Leu	Leu	Met	1090	1095	1100
Val	Ile	Phe	Leu	Leu	Val	Ala	Asn	Ile	Leu	Leu	Asn	Leu	Leu	Ile		1105	1110	1115
Ala	Ile	Phe	Asn	Asn	Ile	Tyr	Asn	Asp	Ser	Ile	Glu	Lys	Ser	Lys	Glu	1125	1130	1135
Ile	Trp	Leu	Phe	Gln	Arg	Tyr	Gln	Gln	Leu	Met	Glu	Tyr	His	Asp	Ser	1140	1145	1150
Pro	Phe	Leu	Pro	Pro	Pro	Phe	Ser	Ile	Phe	Ala	His	Val	Tyr	His	Phe	1155	1160	1165
Ile	Asp	Tyr	Leu	Tyr	Asn	Leu	Arg	Arg	Pro	Asp	Thr	Lys	Arg	Phe	Arg	1170	1175	1180
Ser	Glu	His	Ser	Ile	Lys	Leu	Ser	Val	Thr	Glu	Asp	Glu	Met	Lys	Arg	1185	1190	1195
Ile	Gln	Asp	Phe	Glu	Glu	Asp	Cys	Ile	Asp	Thr	Leu	Thr	Arg	Ile	Arg	1205	1210	1215
Lys	Leu	Lys	Leu	Asn	Thr	Lys	Glu	Pro	Leu	Ser	Val	Thr	Asp	Leu	Thr	1220	1225	1230
Glu	Leu	Thr	Cys	Gln	Arg	Val	His	Asp	Leu	Met	Gln	Glu	Asn	Phe	Leu	1235	1240	1245
Leu	Lys	Ser	Arg	Val	Tyr	Asp	Ile	Glu	Thr	Lys	Ile	Asp	His	Ile	Ser	1250	1255	1260
Asn	Ser	Ser	Asp	Glu	Val	Val	Gln	Ile	Leu	Lys	Asn	Lys	Lys	Leu	Ser	1265	1270	1275
Gln	Asn	Phe	Ala	Ala	Ser	Ser	Leu	Ser	Leu	Pro	Asp	Thr	Ser	Ile	Glu	1285	1290	1295
Val	Pro	Lys	Ile	Thr	Lys	Thr	Leu	Ile	Asp	Cys	His	Leu	Ser	Pro	Val	1300	1305	1310
Ser	Ile	Glu	Asp	Arg	Leu	Ala	Thr	Arg	Ser	Pro	Leu	Leu	Ala	Asn	Leu	1315	1320	1325

-25-

Gln Arg Asp His Thr Leu Arg Lys Leu Pro Thr Trp Glu Thr Ser Thr
 1330 1335 1340
 Ala Ser Thr Ser Ser Phe Glu Phe Val Phe Tyr Phe Thr Arg His Glu
 1345 1350 1355 136
 Gly Asn Glu Asn Lys Tyr Glu Phe Lys Lys Leu Glu Lys Gly Gly Phe
 1365 1370 1375
 Trp Arg Asn Asn Tyr Val Ile Ser Trp Arg Leu
 1380 1385

<210> 15
 <211> 1868
 <212> PRT
 <213> C. Elegans

<400> 15
 Met Asn Leu Cys Tyr Arg Arg His Arg Tyr Ala Ser Ser Pro Glu Val
 1 5 10 15
 Trp Cys Thr Met Glu Ser Asp Glu Leu Gly Val Thr Arg Tyr Leu Gln
 20 25 30
 Ser Lys Gly Gly Asp Gln Val Pro Thr Ser Thr Thr Gly Gly
 35 40 45
 Ala Gly Gly Asp Gly Asn Ala Val Pro Thr Thr Ser Gln Ala Gln Ala
 50 55 60
 Gln Thr Phe Asn Ser Gly Arg Gln Thr Thr Gly Met Ser Ser Gly Asp
 65 70 75 80
 Arg Leu Asn Glu Asp Val Ser Ala Thr Ala Asn Ser Ala Gln Leu Val
 85 90 95
 Leu Pro Thr Pro Leu Phe Asn Gln Met Arg Phe Thr Glu Ser Asn Met
 100 105 110
 Ser Leu Asn Arg His Asn Trp Val Arg Glu Thr Phe Thr Arg Arg Glu
 115 120 125
 Cys Ser Arg Phe Ile Ala Ser Ser Arg Asp Leu His Lys Cys Gly Cys
 130 135 140
 Gly Arg Thr Arg Asp Ala His Arg Asn Ile Pro Glu Leu Thr Ser Glu
 145 150 155 160
 Phe Leu Arg Gln Lys Arg Ser Val Ala Ala Leu Glu Gln Gln Arg Ser
 165 170 175
 Ile Ser Asn Val Asn Asp Asp Ile Asn Thr Gln Asn Met Tyr Thr Lys
 180 185 190
 Arg Gly Ala Asn Glu Lys Trp Ser Leu Arg Lys His Thr Val Ser Leu
 195 200 205
 Ala Thr Asn Ala Phe Gly Gln Val Glu Phe Gln Gly Gly Pro His Pro
 210 215 220
 Tyr Lys Ala Gln Tyr Val Arg Val Asn Phe Asp Thr Glu Pro Ala Tyr
 225 230 235 240
 Ile Met Ser Leu Phe Glu His Val Trp Gln Ile Ser Pro Pro Arg Leu
 245 250 255
 Ile Ile Thr Val His Gly Gly Thr Ser Asn Phe Asp Leu Gln Pro Lys
 260 265 270
 Leu Ala Arg Val Phe Arg Lys Gly Leu Leu Lys Ala Ala Ser Thr Thr
 275 280 285
 Gly Ala Trp Ile Ile Thr Ser Gly Cys Asp Thr Gly Val Val Lys His
 290 295 300
 Val Ala Ala Ala Leu Glu Gly Ala Gln Ser Ala Gln Arg Asn Lys Ile
 305 310 315 320
 Val Cys Ile Gly Ile Ala Pro Trp Gly Leu Lys Lys Arg Glu Asp
 325 330 335
 Phe Ile Gly Gln Asp Lys Thr Val Pro Tyr Tyr Pro Ser Ser Ser Lys
 340 345 350
 Gly Arg Phe Thr Gly Leu Asn Asn Arg His Ser Tyr Phe Leu Leu Val

355	360	365
Asp Asn Gly Thr Val Gly	Arg Tyr Gly Ala Glu Val	Ile Leu Arg Lys
370	375	380
Arg Leu Glu Met Tyr Ile	Ser Gln Lys Gln Lys Ile	Phe Gly Gly Thr
385	390	395
Arg Ser Val Pro Val Val	Cys Val Val Leu Glu Gly	Gly Ser Cys Thr
405	410	415
Ile Arg Ser Val Leu Asp	Tyr Val Thr Asn Val Pro	Arg Val Pro Val
420	425	430
Val Val Cys Asp Gly Ser	Gly Arg Ala Ala Asp	Leu Leu Ala Phe Ala
435	440	445
His Gln Asn Val Thr Glu	Asp Gly Leu Leu Pro	Asp Asp Ile Arg Arg
450	455	460
Gln Val Leu Leu Leu Val	Glu Thr Thr Phe Gly	Cys Ser Glu Ala Ala
465	470	475
Ala His Arg Leu Leu His	Glu Leu Thr Val Cys	Ala Gln His Lys Asn
485	490	495
Leu Leu Thr Ile Phe Arg	Leu Gly Glu Gln Gly	Glu His Asp Val Asp
500	505	510
His Ala Ile Leu Thr Ala	Leu Leu Lys Gly Gln	Asn Leu Ser Ala Ala
515	520	525
Asp Gln Leu Ala Leu Ala	Leu Ala Trp Asn Arg	Val Asp Ile Ala Arg
530	535	540
Ser Asp Val Phe Ala Met	Gly His Glu Trp Pro	Gln Ala Ala Leu His
545	550	555
Asn Ala Met Met Glu Ala	Leu Ile His Asp Arg	Val Asp Phe Val Arg
565	570	575
Leu Leu Leu Glu Gln Gly	Ile Asn Met Gln Lys	Phe Leu Thr Ile Ser
580	585	590
Arg Leu Asp Glu Leu Tyr	Asn Thr Asp Lys Gly	Pro Pro Asn Thr Leu
595	600	605
Phe Tyr Ile Val Arg Asp	Val Val Arg Val Arg	Gln Gly Tyr Arg Phe
610	615	620
Lys Leu Pro Asp Ile Gly	Leu Val Ile Glu Lys	Leu Met Gly Asn Ser
625	630	635
Tyr Gln Cys Ser Tyr Thr	Thr Ser Glu Phe Arg	Asp Lys Tyr Lys Gln
645	650	655
Arg Met Lys Arg Val Lys	His Ala Gln Lys Lys	Ala Met Gly Val Phe
660	665	670
Ser Ser Arg Pro Ser Arg	Thr Gly Ser Gly Ile	Ala Ser Arg Gln Ser
675	680	685
Thr Glu Gly Met Gly Gly	Val Gly Gly Gly Ser	Ser Val Ala Gly Val
690	695	700
Phe Gly Asn Ser Phe Gly	Asn Gln Asp Pro Pro	Leu Asp Pro His Val
705	710	715
Asn Arg Ser Ala Leu Ser	Gly Ser Arg Ala Leu	Ser Asn His Ile Leu
725	730	735
Trp Arg Ser Ala Phe Arg	Gly Asn Phe Pro Ala	Asn Pro Met Arg Pro
740	745	750
Pro Asn Leu Gly Asp Ser	Arg Asp Cys Gly Ser	Glu Phe Asp Glu Glu
755	760	765
Leu Ser Leu Thr Ser Ala	Ser Asp Gly Ser Gln	Thr Glu Pro Asp Phe
770	775	780
Arg Tyr Pro Tyr Ser Glu	Leu Met Ile Trp Ala	Val Leu Thr Lys Arg
785	790	795
Gln Asp Met Ala Met Cys	Met Trp Gln His Gly	Glu Glu Ala Met Ala
805	810	815
Lys Ala Leu Val Ala Cys	Arg Leu Tyr Lys Ser	Leu Ala Thr Glu Ala
820	825	830
Ala Glu Asp Tyr Leu Glu	Val Glu Ile Cys Glu	Glu Leu Lys Lys Tyr

[illegible]

-28-

1315					1320					1325					
Asp	Glu	Ile	Asp	Thr	Cys	Gly	Asp	Glu	Ala	Trp	Asp	Gln	His	Leu	Glu
1330					1335					1340					
Asn	Gly	Gly	Pro	Val	Ile	Leu	Gly	Asn	Gly	Thr	Thr	Gly	Leu	Ser	Cys
1345					1350					1355					
Val	Pro	Gly	Tyr	Trp	Ile	Pro	Pro	Leu	Leu	Met	Thr	Phe	Phe	Leu	Leu
1365					1370					1375					
Ile	Ala	Asn	Ile	Leu	Leu	Met	Ser	Met	Leu	Ile	Ala	Ile	Phe	Asn	His
1380					1385					1390					
Ile	Phe	Asp	Ala	Thr	Asp	Glu	Met	Ser	Gln	Gln	Ile	Trp	Leu	Phe	Gln
1395					1400					1405					
Arg	Tyr	Lys	Gln	Val	Met	Glu	Tyr	Glu	Ser	Thr	Pro	Phe	Leu	Pro	Pro
1410					1415					1420					
Pro	Leu	Thr	Pro	Leu	Tyr	His	Gly	Val	Leu	Ile	Leu	Gln	Phe	Val	Arg
1425					1430					1435					
Thr	Arg	Leu	Ser	Cys	Ser	Lys	Ser	Gln	Glu	Arg	Asn	Pro	Ile	Leu	Leu
1445					1450					1455					
Leu	Lys	Ile	Ala	Glu	Leu	Phe	Leu	Asp	Asn	Asp	Gln	Ile	Glu	Lys	Leu
1460					1465					1470					
His	Asp	Phe	Glu	Glu	Asp	Cys	Met	Glu	Asp	Leu	Ala	Arg	Gln	Lys	Leu
1475					1480					1485					
Asn	Glu	Lys	Asn	Thr	Ser	Asn	Glu	Gln	Arg	Ile	Leu	Arg	Ala	Asp	Ile
1490					1495					1500					
Arg	Thr	Asp	Gln	Ile	Leu	Asn	Arg	Leu	Ile	Asp	Leu	Gln	Ala	Lys	Glu
1505					1510					1515					
Ser	Met	Gly	Arg	Asp	Val	Ile	Asn	Asp	Val	Glu	Ser	Arg	Leu	Ala	Ser
1525					1530					1535					
Val	Glu	Lys	Ala	Gln	Asn	Glu	Ile	Leu	Glu	Cys	Val	Arg	Ala	Leu	Leu
1540					1545					1550					
Asn	Gln	Asn	Asn	Ala	Pro	Thr	Ala	Ile	Gly	Arg	Cys	Phe	Ser	Pro	Ser
1555					1560					1565					
Pro	Asp	Pro	Leu	Val	Glu	Thr	Ala	Asn	Gly	Thr	Pro	Gly	Pro	Leu	Leu
1570					1575					1580					
Leu	Lys	Leu	Pro	Gly	Thr	Asp	Pro	Ile	Leu	Glu	Glu	Lys	Asp	His	Asp
1585					1590					1595					
Ser	Gly	Glu	Asn	Ser	Asn	Ser	Leu	Pro	Pro	Gly	Arg	Ile	Arg	Arg	Asn
1605					1610					1615					
Arg	Thr	Ala	Thr	Ile	Cys	Gly	Gly	Tyr	Val	Ser	Glu	Glu	Arg	Asn	Met
1620					1625					1630					
Met	Leu	Leu	Ser	Pro	Lys	Pro	Ser	Asp	Val	Ser	Gly	Ile	Pro	Gln	Gln
1635					1640					1645					
Arg	Leu	Met	Ser	Val	Thr	Ser	Met	Asp	Pro	Leu	Pro	Leu	Pro	Leu	Ala
1650					1655					1660					
Lys	Leu	Ser	Thr	Met	Ser	Ile	Arg	Arg	Arg	His	Glu	Glu	Tyr	Thr	Ser
1665					1670					1675					
Ile	Thr	Asp	Ser	Ile	Ala	Ile	Arg	His	Pro	Glu	Arg	Arg	Ile	Arg	Asn
1685					1690					1695					
Asn	Arg	Ser	Asn	Ser	Ser	Glu	His	Asp	Glu	Ser	Ala	Val	Asp	Ser	Glu
1700					1705					1710					
Gly	Gly	Gly	Asn	Val	Thr	Ser	Ser	Pro	Arg	Lys	Arg	Ser	Thr	Arg	Asp
1715					1720					1725					
Leu	Arg	Met	Thr	Pro	Ser	Ser	Gln	Val	Glu	Glu	Ser	Thr	Ser	Arg	Asp
1730					1735					1740					
Gln	Ile	Phe	Glu	Ile	Asp	His	Pro	Glu	His	Glu	Glu	Asp	Glu	Ala	Gln
1745					1750					1755					
Ala	Asp	Cys	Glu	Leu	Thr	Asp	Val	Ile	Thr	Glu	Glu	Glu	Asp	Glu	Glu
1765					1770					1775					
Glu	Asp	Asp	Glu	Glu	Asp	Asp	Ser	His	Glu	Arg	His	His	Ile	His	Pro
1780					1785					1790					
Arg	Arg	Lys	Ser	Ser	Arg	Gln	Asn	Arg	Gln	Pro	Ser	His	Thr	Leu	Glu

-29-

1795	1800	1805
Thr Asp Leu Ser Glu Gly	Glu Glu Val Asp Pro	Leu Asp Val Leu Lys
1810	1815	1820
Met Lys Glu Leu Pro Ile	Ile His Gln Ile Leu	Asn Glu Glu Glu Gln
1825	1830	1835
Ala Gly Ala Pro His Ser	Thr Pro Val Ile Ala	Ser Pro Ser Ser Ser
1845	1850	1855
Arg Ala Asp Leu Thr Ser	Gln Lys Cys Ser Asp	Val
1860	1865	

<210> 16
 <211> 489
 <212> DNA
 <213> Mus Musculus

<400> 16	
ccctgaaaga ctcgacttct gctgctagcg ctggagctga gttagttttg agaaggtttc	60
ccgggggctgt ccttggttcgg tggcccggtgc caccgcctcc ggagacgctt tccgatagat	120
ggctgcaggc cgcgagggtg gaggaggagc cgctgccctt ccggagtccg ccccgtagag	180
agaatgtccc agaaatcctg gatagagagc actttgacca agaggggagtg tgtatatatt	240
ataccaagct ccaaagaccc tcacagatgt cttccaggat gtcagatttg tcagcaactt	300
gtcagatggt tctgtggtcg tttggtcaag caacatgcat gctttactgc aagtcttgcc	360
atgaaatact cagatgtgaa attgggtgaa cactttaacc aggcaataga agaatggtct	420
gtggaaaagc acacggagca gagcccaaca gatgcttatg gagtcatcaa ttttcaaggg	480
ggttctcat	489

<210> 17
 <211> 102
 <212> PRT
 <213> Mus Musculus

<400> 17	
Met Ser Gln Lys Ser Trp Ile Glu Ser Thr Leu Thr Lys Arg Glu Cys	
1 5 10 15	
Val Tyr Ile Ile Pro Ser Ser Lys Asp Pro His Arg Cys Leu Pro Gly	
20 25 30	
Cys Gln Ile Cys Gln Gln Leu Val Arg Cys Phe Cys Gly Arg Leu Val	
35 40 45	
Lys Gln His Ala Cys Phe Thr Ala Ser Leu Ala Met Lys Tyr Ser Asp	
50 55 60	
Val Lys Leu Gly Glu His Phe Asn Gln Ala Ile Glu Glu Trp Ser Val	
65 70 75 80	
Glu Lys His Thr Glu Gln Ser Pro Thr Asp Ala Tyr Gly Val Ile Asn	
85 90 95	
Phe Gln Gly Gly Ser His	
100	

<210> 18
 <211> 410
 <212> DNA
 <213> Homo Sapiens

<220>
 <221> unsure
 <222> (6)...(6)
 <221> unsure
 <222> (58)...(58)
 <221> unsure

<222> (89)...(89)

<221> unsure

<222> (406)...(406)

<400> 18

gccgcnggag	cctgagcgga	gggtgtgcg	agcctcgcca	gcgggggccc	cggtctgngc	60
cattgcctca	ctgagccagc	gcctgcctnc	tacctcgccg	acagctggaa	ccagtgcgac	120
ctagtggctc	tcacctgctt	cctcctgggc	gtgggctgcc	ggctgacccc	gggtttgtac	180
cacctgggccc	gcactgtcct	ctgcatcgac	ttcatggttt	tcacgggtcg	gctgcttcac	240
atcttcacgg	tcaacaaaca	gctggggccc	aagatcgta	tcgtgagcaa	gatgatgaag	300
gacgtgttct	tcttcctctt	cttcctcggc	gtgtggctgg	tagctatggg	ttggggccacg	360
gaggggttcc	tgaggccacg	ggacagtgc	ttcccaagta	tcctgncgcc		410

<210> 19

<211> 131

<212> PRT

<213> Homo Sapiens

<220>

<221> UNSURE

<222> (15)...(15)

<223> UNKNOWN

<221> UNSURE

<222> (25)...(25)

<223> UNKNOWN

<221> UNSURE

<222> (131)...(131)

<223> UNKNOWN

<400> 19

Ala	Glu	Gly	Val	Arg	Ser	Leu	Ala	Ser	Gly	Gly	Pro	Gly	Leu	Xaa	His
1				5					10					15	
Cys	Leu	Thr	Glu	Pro	Ala	Pro	Ala	Xaa	Tyr	Leu	Ala	Asp	Ser	Trp	Asn
			20					25					30		
Gln	Cys	Asp	Leu	Val	Ala	Leu	Thr	Cys	Phe	Leu	Leu	Gly	Val	Gly	Cys
		35					40					45			
Arg	Leu	Thr	Pro	Gly	Leu	Tyr	His	Leu	Gly	Arg	Thr	Val	Leu	Cys	Ile
		50				55					60				
Asp	Phe	Met	Val	Phe	Thr	Val	Arg	Leu	Leu	His	Ile	Phe	Thr	Val	Asn
					70					75					80
Lys	Gln	Leu	Gly	Pro	Lys	Ile	Val	Ile	Val	Ser	Lys	Met	Met	Lys	Asp
				85					90					95	
Val	Phe	Phe	Phe	Leu	Phe	Phe	Leu	Gly	Val	Trp	Leu	Val	Ala	Met	Gly
			100					105					110		
Trp	Ala	Thr	Glu	Gly	Phe	Leu	Arg	Pro	Arg	Asp	Ser	Asp	Phe	Pro	Ser
		115					120					125			
Ile	Leu	Xaa													
		130													

<210> 20

<211> 389

<212> DNA

<213> Homo Sapiens

<400> 20

caaatttttt	gttagtacac	catctcatcc	aaattgcaaa	agtcacatgg	aaactggaac	60
caaagatcaa	gaaactgttt	gctctaaagc	tacagaagga	gataatacag	aatttgagac	120

-31-

atttgttagga	cacagagata	gcatggat	acagagggtt	aaagaaacat	caaacaagat	180
aaaaataacta	tccaataaca	atacttctga	aaacactttg	aaacgagtga	gttctcttgc	240
tggatttact	gactgtcaca	gaacttccat	tcctgttcat	tcaaaacgag	aaaagatcag	300
tagaaggcca	tctaccgaag	acactcatga	agtagattcc	aaagcagctt	taataccggt	360
ttgtagattt	caactaaaca	gatatatat				389

<210> 21
 <211> 415
 <212> DNA
 <213> Homo Sapiens

<400> 21						
atttctagtt	tttcaaattt	gccagtcttt	ttgaatagta	tctccttctt	ttctcatggt	60
ttatatattaa	aactttttta	tgtccatcat	cacttttaa	atacttattt	tgtcatctat	120
aaccaataat	tccactatct	tatcagaaat	caaataccgt	ttatgtaagt	tgactcccat	180
gagttctaaa	ttgccattgt	gaggtcactt	tcggtttaggc	tttaatttgt	tgcaaagttg	240
tgcagctcag	ggtcaggaag	agtcctcca	gaaaggagga	tttgttactg	tgaatctctt	300
tgttaactaa	cctctttccc	cactgaaata	acttttttca	ataacatgat	tttaacaaca	360
taatctctct	atgccagaac	agatatatat	gaatgtaagt	caatattttc	ttgag	415

<210> 22
 <211> 405
 <212> DNA
 <213> Mus Musculus

<400> 22						
ttattatggc	ttatcatgaa	aaaccagtcc	tgcctcctcc	tcttatcatc	ctcagccata	60
tagtttctact	gttttgctgt	gtatgcaaaa	gaagaaagaa	agataagact	tccgatgggc	120
caaaactttt	cttaacagaa	gaagatcaaa	agaaactcca	tgattttgaa	gagcagtgtg	180
ttgagatgta	ctttgatgag	aaagatgaca	aattcaattc	tgggagtga	gagagaatcc	240
gggtcacttt	tgaaagagtg	gagcagatga	gcattcagat	taaagaagtt	ggagatcgtg	300
tcaactacat	aaaaagatca	ttacagtctt	tagattctca	aattgggtcat	ctgcaagatc	360
tctcagccct	aacagtagat	acattgaaaa	cacttacagc	ccaga		405

<210> 23
 <211> 5117
 <212> DNA
 <213> Homo Sapiens

<220>
 <221> unsure
 <222> (2382)... (2382)
 <223> unknown

<221> unsure
 <222> (4664)... (4664)
 <223> unknown

<221> unsure
 <222> (4682)... (4682)
 <223> unknown

<221> unsure
 <222> (4702)... (4702)
 <223> unknown

<221> unsure
 <222> (5038)... (5039)
 <223> unknown

-32-

<221> unsure
 <222> (5056)... (5056)
 <223> unknown

<221> unsure
 <222> (5071)... (5072)

<400> 23

gatggcaaca	tggtgaagaa	tcaatggcta	aagcattagt	tgctgtgaag	atctatcggt	60
caatggcata	tgaagcaaag	cagagtgacc	tggtagatga	tacttcagaa	gaactaaaac	120
agtattccaa	tgattttggg	cagttggccg	ttgaattatt	agaacagtc	ttcagacaag	180
atgaaaccat	ggctatgaaa	ttgctcactt	atgaactgaa	gaactggagt	aattcaacct	240
gccttaagtt	agcagtttct	tcaagactta	gaccttttgt	agctcacacc	tgtacacaaa	300
tggtgttata	tgatatgtgg	atgggaaggc	tgaatatgag	gaaaaattcc	tggtacaagg	360
tcatactaag	catttttagtt	ccacctgcca	tattgctgtt	agagtataaa	actaaggctg	420
aaatgtccca	tatcccacaa	tctcaagatg	ctcatcagat	gacaatggat	gacagcgaaa	480
acaactttca	gaacataaca	gaagagatcc	ccatggaagt	gtttaaagaa	gtacggattt	540
tggaatgtaa	tgaaggaaaag	aatgagatgg	agatacaaat	gaaatcaaaa	aagcttccaa	600
ttacgcgaaa	gttttatgcc	ttttatcatg	caccaattgt	aaaattctgg	tttaacacgt	660
tggcataatt	aggattttctg	atgctttata	cattttgtgt	tcttgtaaaa	atggaacagt	720
taccttcagt	tcaagaatgg	attgttattg	cttatatttt	tacttatgcc	attgagaaaag	780
tccgtgagat	ctttatgtct	gaagctggga	aagtaaacca	gaagattaaa	gtatggttta	840
gtgattactt	caacatcagt	gatacaattg	ccataatttc	tttcttcatt	ggatttgagc	900
taagatttgg	agcaaaatgg	aactttgcaa	atgcatatga	taatcatgtt	tttgtggctg	960
gaagattaat	ttactgtctt	aacataatat	tttggatatg	gcgtttgcta	gattttctag	1020
ctgtaaatca	acaggatcaa	ccttatgtaa	tgatgattgg	aaaaatggtg	gccaatatgt	1080
tctacattgt	agtgtattatg	gctcttgat	tacttagttt	tggtgttccc	agaaaggcaa	1140
tactttatcc	tcatgaagca	ccatcttgg	ctcttgctaa	agatatagtt	tttcacccat	1200
actggatgat	ttttggtgaa	gtttatgcat	acgaaattga	tgtgtgtgca	aatgattctg	1260
ttatccctca	aatctgtggt	cctgggacgt	ggttgactcc	atttcttcaa	gcagtctacc	1320
tctttgtaca	gtataatcatt	atggttaata	ttcttattgc	atttttcaac	aatgtgtatt	1380
tacaagtga	ggcaatttcc	aatattgtat	ggaagtacca	gcgttatcat	tttattatgg	1440
cttatcatga	gaaaccagtt	ctgcctcctc	cacttatcat	tcttagccat	atagtttctc	1500
tgttttgctg	catatgtaag	agaagaaaga	aagataagac	ttccgatgga	ccaaaacttt	1560
tcttaacaga	agaagatcaa	aagaaacttc	atgattttga	agagcagtg	gttgaaatgt	1620
atttcaatga	aaaagatgac	aaatttcatt	ctgggagtga	agagagaatt	cgtgtcactt	1680
ttgaaagagt	ggaacagatg	tgcattcaga	ttaaagaagt	tgagatcg	gtcaactaca	1740
taaaaagatc	attacaatca	ttagattctc	aaattggcca	tttgcaagat	ctttcagccc	1800
tgacggtaga	tacattaaaa	acactcactg	cccagaaagc	gtcggaaagc	agcaaagtcc	1860
ataatgaaat	cacacgagaa	ctgagcattt	ccaaacactt	ggctcaaaac	cttattgatg	1920
atggtcctgt	aagaccttct	gtatggaaaa	agcatggtgt	tgtaaataca	cttagctcct	1980
ctcttcctca	aggtgatctt	gaaagtaata	atccttttca	ttgtaataat	ttaatgaaaag	2040
atgacaaaag	tccccagtg	aatatatttg	gtcaagactt	acctgcagta	ccccagagaa	2100
aagaatttaa	ttttccagag	gctggttctt	cttctggtgc	cttattccca	agtgtgtttt	2160
ccctccaga	actgcgacag	agactacatg	gggtagaact	cttaaaaata	tttaataaaa	2220
atcaaaaatt	aggcagttca	tctactagca	taccacatct	gtcatcccca	ccaaccaa	2280
tttttgtag	tacaccatct	cagccaagtt	gcaaaagcca	cttggaact	ggaaccaaag	2340
atcaagaaac	tgtttgcctt	aaagctacag	aaggagataa	tncagaattt	ggagcatttg	2400
taggacacag	agatagcatg	gatttacaga	ggtttaaga	aacatcaaac	aagataaaaa	2460
tactatccaa	taacaatact	tctgaaaaca	cttgaaaacg	agtgtgttct	cttgctggat	2520
ttactgactg	tcacagaact	tccattcctg	ttcattcaaa	acaagcagaa	aaaatcagta	2580
gaaggccatc	taccgaagac	actcatgaag	tagattccaa	agcagcttta	ataccggatt	2640
ggttacaaga	tagaccatca	aacagagaaa	tgccatctga	agaaggaa	ttaaatggct	2700
tcactttctc	atttaagcca	gctatggata	caaattacta	ttattcagct	gtggaaagaa	2760
ataaactgat	gaggttatca	cagagcattc	catttacacc	tggtgcctca	agaggggagc	2820
ctgtcacagt	gtatcgtttg	gaagagagtt	cacccaacat	actaaataac	agcatgtctt	2880
cttggtcaca	actagccctc	tgtgccaaaa	tacagttttt	aagcaaaagag	gagatgggag	2940
gaggtttacg	aagagctgtc	aaagtacagt	gtacctggtc	agaacatgat	atcctcaaat	3000
cagggcatct	ttatattatc	aaatcttttc	ttccagaggt	ggttaataca	tggtcaagta	3060
tttataaaga	agatacagtt	ctgcatctct	gtctgagaga	aattcaacaa	cagagagcag	3120

-33-

```

cacaaaagct tacgtttgcc tttaatcaaa tgaaacccaa atccatacca tattctccaa 3180
ggttccttga agttttcctg ctgtattgcc attcagcagg acagtgggtt gctgtggaag 3240
aatgtatgac tggagaatth agaaaataca acaataataa tggagatgag attattccaa 3300
ctaatactct ggaagagatc atgctagcct ttagccactg gacttacgaa tatacaagag 3360
gggagttact ggtacttgat ttgcaagggt ttggtgaaaa tttgactgac ccatctgtga 3420
taaaagcaga agaaaagaga tcctgtgata tgggtttttg cccagcaaas ctaggagaag 3480
atgcaattaa aaacttcaga gcaaaacatc actgtaattc ttgctgtaga aagcttaaac 3540
ttccagatct gaagaggaat gattatacgc ctgataaaat tatatttcct caggatgagc 3600
cttcagatth gaatcttcag cctggaaatt ccaccaaaga atcagaatca gctaattctg 3660
ttcgtctgat gttataatat taatattact gaatcattgg ttttgctgc acctcacaga 3720
aatgttactg tgtcactttt ccctcgggag gaaattgttt ggtaatatag aaagggtgat 3780
gcaagttgaa tttgctgact ccagcacagt taaaagggtca atattctttt gacctgatta 3840
atcagtcaga aagtccttat aggatagagc tggcagctga gaaattttta aggtaattga 3900
taattagtat ttgtaacttt ttaaagggtc ctttgtatag cagaggatct catttgactt 3960
tgttttgatg aggggtgatgc cctctcttat gtggtacaat accattaacc aaaggtaggt 4020
gtccatgcag attttatttg cagctgtttt attgccattc aactagggaa atgaagaaat 4080
cacgcagcct tttgggttaa tggcagtcaa aattttcttc agtgtattta gtgtgttcag 4140
tgatgatatc actggttccc aactagatgc ttgttggtcc cgggaaggga aatgacttgt 4200
tctaattcta gggtcacaga ggtatgagaa gcctgaactg aagaccattt tcaagaggga 4260
cgggtatttt gaatcagggt taggctccat atttaaagat agagccagtt ttttttttaa 4320
atagaaccca aattgtgtaa aaatgttaat tgggtttttt aaacattggt ttatcaagtc 4380
actgttaagt agaagaaaag catggtaaac tgatacataa cctaaattat aaaagcagaa 4440
acctaaactc ctggtcaagg gaagttacct tttgaggaaa gttaaagtac ttttttccct 4500
atctgtatct atagcaacaa cccagaactt acaaacttct ccaaagattt tattgattgt 4560
tatatcaaat cagaatgtaa acatgaactc ttgcatatat ttaaaattgt gttggaacat 4620
ttgaacatga atgctgtttg ggtacttaag aaattrattc agtnggatta tcattatgtg 4680
anactggcag attgcagtgc anccttatgc caataaaatg taatttaaca gcccagata 4740
ttgttgataa ttcaacaata acaagaaaag cttttcatct aagttttatg ctttaatttt 4800
ttttcttttt ttttcttttt cttttgtttc cttggtacta attttaattt ttatttgga 4860
gggagcagta taaagcttat ttgtatttag tagtgtatct catagatata gacaaggcaa 4920
gagatgataa gctgttttaa tagtgtttta tattgattgg ggggtggggag aaagaaaaag 4980
tgtattactt aaagatacta tatacgtttt gtatatcatt aaatctttta aagaaatnna 5040
ataaatttat tgtttncaaa aaaaaaaccc nntaaaaaaa aaagggcggc ccctctagag 5100
gatccctcga ggggccc 5117

```

<210> 24
 <211> 1224
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> UNSURE
 <222> (794)...(794)
 <223> UNKNOWN

<400> 24
 Trp Gln His Gly Glu Ser Met Ala Lys Ala Leu Val Ala Cys Lys
 1 5 10 15
 Ile Tyr Arg Ser Met Ala Tyr Glu Ala Lys Gln Ser Asp Leu Val Asp
 20 25 30
 Asp Thr Ser Glu Glu Leu Lys Gln Tyr Ser Asn Asp Phe Gly Gln Leu
 35 40 45
 Ala Val Glu Leu Leu Glu Gln Ser Phe Arg Gln Asp Glu Thr Met Ala
 50 55 60
 Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ser Thr Cys
 65 70 75 80
 Leu Lys Leu Ala Val Ser Ser Arg Leu Arg Pro Phe Val Ala His Thr
 85 90 95

-34-

Cys	Thr	Gln	Met	Leu	Leu	Ser	Asp	Met	Trp	Met	Gly	Arg	Leu	Asn	Met
			100					105					110		
Arg	Lys	Asn	Ser	Trp	Tyr	Lys	Val	Ile	Leu	Ser	Ile	Leu	Val	Pro	Pro
		115					120					125			
Ala	Ile	Leu	Leu	Leu	Glu	Tyr	Lys	Thr	Lys	Ala	Glu	Met	Ser	His	Ile
		130				135					140				
Pro	Gln	Ser	Gln	Asp	Ala	His	Gln	Met	Thr	Met	Asp	Asp	Ser	Glu	Asn
145					150					155					160
Asn	Phe	Gln	Asn	Ile	Thr	Glu	Glu	Ile	Pro	Met	Glu	Val	Phe	Lys	Glu
				165					170					175	
Val	Arg	Ile	Leu	Asp	Ser	Asn	Glu	Gly	Lys	Asn	Glu	Met	Glu	Ile	Gln
		180						185					190		
Met	Lys	Ser	Lys	Lys	Leu	Pro	Ile	Thr	Arg	Lys	Phe	Tyr	Ala	Phe	Tyr
		195					200					205			
His	Ala	Pro	Ile	Val	Lys	Phe	Trp	Phe	Asn	Thr	Leu	Ala	Tyr	Leu	Gly
		210				215					220				
Phe	Leu	Met	Leu	Tyr	Thr	Phe	Val	Val	Leu	Val	Gln	Met	Glu	Gln	Leu
225					230					235					240
Pro	Ser	Val	Gln	Glu	Trp	Ile	Val	Ile	Ala	Tyr	Ile	Phe	Thr	Tyr	Ala
				245					250					255	
Ile	Glu	Lys	Val	Arg	Glu	Ile	Phe	Met	Ser	Glu	Ala	Gly	Lys	Val	Asn
			260				265						270		
Gln	Lys	Ile	Lys	Val	Trp	Phe	Ser	Asp	Tyr	Phe	Asn	Ile	Ser	Asp	Thr
		275					280					285			
Ile	Ala	Ile	Ile	Ser	Phe	Phe	Ile	Gly	Phe	Gly	Leu	Arg	Phe	Gly	Ala
		290				295					300				
Lys	Trp	Asn	Phe	Ala	Asn	Ala	Tyr	Asp	Asn	His	Val	Phe	Val	Ala	Gly
305					310				315						320
Arg	Leu	Ile	Tyr	Cys	Leu	Asn	Ile	Ile	Phe	Trp	Tyr	Val	Arg	Leu	Leu
				325					330					335	
Asp	Phe	Leu	Ala	Val	Asn	Gln	Gln	Ala	Gly	Pro	Tyr	Val	Met	Met	Ile
			340					345					350		
Gly	Lys	Met	Val	Ala	Asn	Met	Phe	Tyr	Ile	Val	Val	Ile	Met	Ala	Leu
		355					360					365			
Val	Leu	Leu	Ser	Phe	Gly	Val	Pro	Arg	Lys	Ala	Ile	Leu	Tyr	Pro	His
		370				375					380				
Glu	Ala	Pro	Ser	Trp	Thr	Leu	Ala	Lys	Asp	Ile	Val	Phe	His	Pro	Tyr
385					390					395					400
Trp	Met	Ile	Phe	Gly	Glu	Val	Tyr	Ala	Tyr	Glu	Ile	Asp	Val	Cys	Ala
				405					410					415	
Asn	Asp	Ser	Val	Ile	Pro	Gln	Ile	Cys	Gly	Pro	Gly	Thr	Trp	Leu	Thr
			420					425					430		
Pro	Phe	Leu	Gln	Ala	Val	Tyr	Leu	Phe	Val	Gln	Tyr	Ile	Ile	Met	Val
		435					440					445			
Asn	Leu	Leu	Ile	Ala	Phe	Phe	Asn	Asn	Val	Tyr	Leu	Gln	Val	Lys	Ala
		450				455					460				
Ile	Ser	Asn	Ile	Val	Trp	Lys	Tyr	Gln	Arg	Tyr	His	Phe	Ile	Met	Ala
465					470					475					480
Tyr	His	Glu	Lys	Pro	Val	Leu	Pro	Pro	Pro	Leu	Ile	Ile	Leu	Ser	His
				485					490					495	
Ile	Val	Ser	Leu	Phe	Cys	Cys	Ile	Cys	Lys	Arg	Arg	Lys	Lys	Asp	Lys
			500					505					510		
Thr	Ser	Asp	Gly	Pro	Lys	Leu	Phe	Leu	Thr	Glu	Glu	Asp	Gln	Lys	Lys
		515					520					525			
Leu	His	Asp	Phe	Glu	Glu	Gln	Cys	Val	Glu	Met	Tyr	Phe	Asn	Glu	Lys
		530				535					540				
Asp	Asp	Lys	Phe	His	Ser	Gly	Ser	Glu	Glu	Arg	Ile	Arg	Val	Thr	Phe
545					550					555					560
Glu	Arg	Val	Glu	Gln	Met	Cys	Ile	Gln	Ile	Lys	Glu	Val	Gly	Asp	Arg
				565					570					575	

-35-

Val	Asn	Tyr	Ile	Lys	Arg	Ser	Leu	Gln	Ser	Leu	Asp	Ser	Gln	Ile	Gly
			580					585					590		
His	Leu	Gln	Asp	Leu	Ser	Ala	Leu	Thr	Val	Asp	Thr	Leu	Lys	Thr	Leu
		595					600					605			
Thr	Ala	Gln	Lys	Ala	Ser	Glu	Ala	Ser	Lys	Val	His	Asn	Glu	Ile	Thr
	610					615				620					
Arg	Glu	Leu	Ser	Ile	Ser	Lys	His	Leu	Ala	Gln	Asn	Leu	Ile	Asp	Asp
625					630				635					640	
Gly	Pro	Val	Arg	Pro	Ser	Val	Trp	Lys	Lys	His	Gly	Val	Val	Asn	Thr
			645					650						655	
Leu	Ser	Ser	Ser	Leu	Pro	Gln	Gly	Asp	Leu	Glu	Ser	Asn	Asn	Pro	Phe
		660						665					670		
His	Cys	Asn	Ile	Leu	Met	Lys	Asp	Asp	Lys	Asp	Pro	Gln	Cys	Asn	Ile
	675					680						685			
Phe	Gly	Gln	Asp	Leu	Pro	Ala	Val	Pro	Gln	Arg	Lys	Glu	Phe	Asn	Phe
	690					695				700					
Pro	Glu	Ala	Gly	Ser	Ser	Ser	Gly	Ala	Leu	Phe	Pro	Ser	Ala	Val	Ser
705					710					715					720
Pro	Pro	Glu	Leu	Arg	Gln	Arg	Leu	His	Gly	Val	Glu	Leu	Leu	Lys	Ile
			725						730					735	
Phe	Asn	Lys	Asn	Gln	Lys	Leu	Gly	Ser	Ser	Ser	Thr	Ser	Ile	Pro	His
	740						745						750		
Leu	Ser	Ser	Pro	Pro	Thr	Lys	Phe	Phe	Val	Ser	Thr	Pro	Ser	Gln	Pro
	755					760						765			
Ser	Cys	Lys	Ser	His	Leu	Glu	Thr	Gly	Thr	Lys	Asp	Gln	Glu	Thr	Val
	770					775					780				
Cys	Ser	Lys	Ala	Thr	Glu	Gly	Asp	Asn	Xaa	Glu	Phe	Gly	Ala	Phe	Val
785					790				795					800	
Gly	His	Arg	Asp	Ser	Met	Asp	Leu	Gln	Arg	Phe	Lys	Glu	Thr	Ser	Asn
			805					810						815	
Lys	Ile	Lys	Ile	Leu	Ser	Asn	Asn	Asn	Thr	Ser	Glu	Asn	Thr	Leu	Lys
			820					825					830		
Arg	Val	Ser	Ser	Leu	Ala	Gly	Phe	Thr	Asp	Cys	His	Arg	Thr	Ser	Ile
	835					840						845			
Pro	Val	His	Ser	Lys	Gln	Ala	Glu	Lys	Ile	Ser	Arg	Arg	Pro	Ser	Thr
	850					855					860				
Glu	Asp	Thr	His	Glu	Val	Asp	Ser	Lys	Ala	Ala	Leu	Ile	Pro	Asp	Trp
865					870				875					880	
Leu	Gln	Asp	Arg	Pro	Ser	Asn	Arg	Glu	Met	Pro	Ser	Glu	Glu	Gly	Thr
			885					890						895	
Leu	Asn	Gly	Leu	Thr	Ser	Pro	Phe	Lys	Pro	Ala	Met	Asp	Thr	Asn	Tyr
		900						905					910		
Tyr	Tyr	Ser	Ala	Val	Glu	Arg	Asn	Asn	Leu	Met	Arg	Leu	Ser	Gln	Ser
	915						920					925			
Ile	Pro	Phe	Thr	Pro	Val	Pro	Pro	Arg	Gly	Glu	Pro	Val	Thr	Val	Tyr
	930					935					940				
Arg	Leu	Glu	Glu	Ser	Ser	Pro	Asn	Ile	Leu	Asn	Asn	Ser	Met	Ser	Ser
945					950					955				960	
Trp	Ser	Gln	Leu	Gly	Leu	Cys	Ala	Lys	Ile	Glu	Phe	Leu	Ser	Lys	Glu
			965					970						975	
Glu	Met	Gly	Gly	Gly	Leu	Arg	Arg	Ala	Val	Lys	Val	Gln	Cys	Thr	Trp
		980						985					990		
Ser	Glu	His	Asp	Ile	Leu	Lys	Ser	Gly	His	Leu	Tyr	Ile	Ile	Lys	Ser
	995						1000					1005			
Phe	Leu	Pro	Glu	Val	Val	Asn	Thr	Trp	Ser	Ser	Ile	Tyr	Lys	Glu	Asp
	1010					1015						1020			
Thr	Val	Leu	His	Leu	Cys	Leu	Arg	Glu	Ile	Gln	Gln	Gln	Arg	Ala	Ala
1025					1030					1035				104	
Gln	Lys	Leu	Thr	Phe	Ala	Phe	Asn	Gln	Met	Lys	Pro	Lys	Ser	Ile	Pro
				1045					1050					1055	

-36-

Tyr Ser Pro Arg Phe Leu Glu Val Phe Leu Leu Tyr Cys His Ser Ala
 1060 1065 1070
 Gly Gln Trp Phe Ala Val Glu Glu Cys Met Thr Gly Glu Phe Arg Lys
 1075 1080 1085
 Tyr Asn Asn Asn Asn Gly Asp Glu Ile Ile Pro Thr Asn Thr Leu Glu
 1090 1095 1100
 Glu Ile Met Leu Ala Phe Ser His Trp Thr Tyr Glu Tyr Thr Arg Gly
 1105 1110 1115 112
 Glu Leu Leu Val Leu Asp Leu Gln Gly Val Gly Glu Asn Leu Thr Asp
 1125 1130 1135
 Pro Ser Val Ile Lys Ala Glu Glu Lys Arg Ser Cys Asp Met Val Phe
 1140 1145 1150
 Gly Pro Ala Asn Leu Gly Glu Asp Ala Ile Lys Asn Phe Arg Ala Lys
 1155 1160 1165
 His His Cys Asn Ser Cys Cys Arg Lys Leu Lys Leu Pro Asp Leu Lys
 1170 1175 1180
 Arg Asn Asp Tyr Thr Pro Asp Lys Ile Ile Phe Pro Gln Asp Glu Pro
 1185 1190 1195 120
 Ser Asp Leu Asn Leu Gln Pro Gly Asn Ser Thr Lys Glu Ser Glu Ser
 1205 1210 1215
 Ala Asn Ser Val Arg Leu Met Leu
 1220

<210> 25

<211> 2180

<212> DNA

<213> Homo Sapiens

<400> 25

tcgaggccaa	gaattcggca	cgagggcctc	gggcaggccc	cctggagcga	cctgctttctt	60
tgggcactgt	tgctgaacag	ggcacagatg	gccatgtact	tctgggagat	gggttccaat	120
gcagtttctt	cagctcttgg	ggcctgtttg	ctgctccggg	tgatggcacg	cctgggagcct	180
gacgctgagg	aggcagcacg	gaggaaaagac	ctggcggttca	agtttgaggg	gatgggcgtt	240
gacctctttg	gcgagtgcta	tcgcagcagt	gaggtgaggg	ctgccgcctt	cctcctccgt	300
cgctgcccgc	tctgggggga	tgccacttgc	ctccagctgg	ccatgcaagc	tgacgcccgt	360
gccttctttg	cccaggatgg	ggtacagtct	ctgctgacac	agaagtgggtg	gggagatatg	420
gccagcacta	cacccatctg	ggccctgggt	ctgcgcttct	tttgccttcc	actcatctac	480
acccgcctca	tcaccttcag	gaaatcagaa	gaggagccca	cacgggagga	gctagagttt	540
gacatggata	gtgtcattaa	tggggaaggg	cctgtcggga	cgccggaccc	agccgagaag	600
acgccgctgg	gggtcccgcg	ccagtcgggc	cgctccgggtt	gctgcggggg	ccgctgcggg	660
gggcgcgggt	gcctacgccc	ctggttccac	ttctggggcg	cgccgggtgac	catcttcatg	720
ggcaacgtgg	tcagctacct	gctgttctct	ctgcttttct	cgccgggtgct	gctcgtggat	780
ttccagccgg	cgccgcccgg	ctccctggag	ctgctgctct	atttctgggc	tttcacgctg	840
ctgtgcgagg	aactgcgcca	gggcctgagc	ggaggcgggg	gcagcctcgc	cagcgggggc	900
cccgggcctg	gccatgcctc	actgagccag	cgccctgcgc	tctacctcgc	cgacagctgg	960
aaccagtgcg	acctagtggc	tctcacctgc	ttcctcctgg	gcgtgggctg	ccggtgacc	1020
ccgggtttgt	accacctggg	cgcactatgc	ctctgcctgc	acttcatggt	tttcacggtg	1080
cggtgcttcc	acatcttcac	ggtaacaaaa	cagctggggc	ccaagatcgt	catcgtgagc	1140
aagatgatga	aggacgtgtt	cttcttcttc	ttcttctctg	gcgtgtgggt	ggtagcctat	1200
ggcgtggcca	cggaggggct	cctgaggcca	cgggacagtg	acttcccaag	tatcctgcgc	1260
cgcgtcttct	accgtcccta	cctgcagatc	ttcgggcaga	ttccccagga	ggacatggac	1320
gtggccctca	tggagcacag	caactgctcg	tcggagcccg	gcttctgggc	acaccctcct	1380
ggggcccagg	cgggcacctg	cgtctcccag	tatgccaaact	ggctgggtgg	gctgctcttc	1440
gtcatcttcc	tgctcgtggc	caacatcctg	ctgggtcaact	tgctcattgc	catgttcagt	1500
tacacattcg	gcaaagtaca	gggcaacagc	gatctctact	ggaaggcgca	gcgttacccg	1560
ctcatccggg	aattccactc	tcggcccgcg	ctggccccgc	cctttatcgt	catctcccac	1620
ttgcgcctcc	tgctcaggca	attgtgcagg	cgacccsgga	gccccagcc	gtcctccccg	1680
gccctcgagc	atttccgggt	ttacctttct	aagggaagccg	agcgggaagct	gctaacgtgg	1740
gaatcggtgc	ataaggagaa	ctttctgctg	gcacgcgcta	gggacaagcg	ggagagcgac	1800
tccgagmgtc	tgaagcgcac	gtcccagaag	gtggacttgg	cactgaaaca	gctgggacac	1860

-37-

```

atccgcgagt acgaacagcg cctgaaagtg ctggagcggg aggtccagca gtgtacctcg 1920
gccccgcac ctggtggcct tgccttgag gtgagcccca tgtccatctg ggccactgtc 1980
aggaccacct ttgggagtgat catccttaca aaccacagca tgcccggctc ctcccagaac 2040
cagtcccagc ctgggaggat caaggcctgg atcccrggcc gttatccatc tggaggctgc 2100
agggtccttg gggtaacagg gaccacagac ccctcaccac tcacagattc ctcacactgg 2160
ggaaataaag ccatttcaga 2180

```

<210> 26
 <211> 725
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> UNSURE
 <222> (553)...(553)
 <223> UNKNOWN

<221> UNSURE
 <222> (603)...(603)
 <223> UNKNOWN

<400> 26

Ser	Arg	Pro	Arg	Ile	Arg	His	Glu	Gly	Leu	Gly	Gln	Ala	Pro	Trp	Ser	1	5	10	15
Asp	Leu	Leu	Leu	Trp	Ala	Leu	Leu	Leu	Asn	Arg	Ala	Gln	Met	Ala	Met	20	25	30	
Tyr	Phe	Trp	Glu	Met	Gly	Ser	Asn	Ala	Val	Ser	Ser	Ala	Leu	Gly	Ala	35	40	45	
Cys	Leu	Leu	Leu	Arg	Val	Met	Ala	Arg	Leu	Glu	Pro	Asp	Ala	Glu	Glu	50	55	60	
Ala	Ala	Arg	Arg	Lys	Asp	Leu	Ala	Phe	Lys	Phe	Glu	Gly	Met	Gly	Val	65	70	75	80
Asp	Leu	Phe	Gly	Glu	Cys	Tyr	Arg	Ser	Ser	Glu	Val	Arg	Ala	Ala	Arg	85	90	95	
Leu	Leu	Leu	Arg	Arg	Cys	Pro	Leu	Trp	Gly	Asp	Ala	Thr	Cys	Leu	Gln	100	105	110	
Leu	Ala	Met	Gln	Ala	Asp	Ala	Arg	Ala	Phe	Phe	Ala	Gln	Asp	Gly	Val	115	120	125	
Gln	Ser	Leu	Leu	Thr	Gln	Lys	Trp	Trp	Gly	Asp	Met	Ala	Ser	Thr	Thr	130	135	140	
Pro	Ile	Trp	Ala	Leu	Val	Leu	Ala	Phe	Phe	Cys	Pro	Pro	Leu	Ile	Tyr	145	150	155	160
Thr	Arg	Leu	Ile	Thr	Phe	Arg	Lys	Ser	Glu	Glu	Glu	Pro	Thr	Arg	Glu	165	170	175	
Glu	Leu	Glu	Phe	Asp	Met	Asp	Ser	Val	Ile	Asn	Gly	Glu	Gly	Pro	Val	180	185	190	
Gly	Thr	Ala	Asp	Pro	Ala	Glu	Lys	Thr	Pro	Leu	Gly	Val	Pro	Arg	Gln	195	200	205	
Ser	Gly	Arg	Pro	Gly	Cys	Cys	Gly	Gly	Arg	Cys	Gly	Gly	Arg	Arg	Cys	210	215	220	
Leu	Arg	Arg	Trp	Phe	His	Phe	Trp	Gly	Ala	Pro	Val	Thr	Ile	Phe	Met	225	230	235	240
Gly	Asn	Val	Val	Ser	Tyr	Leu	Leu	Phe	Leu	Leu	Leu	Phe	Ser	Arg	Val	245	250	255	
Leu	Leu	Val	Asp	Phe	Gln	Pro	Ala	Pro	Pro	Gly	Ser	Leu	Glu	Leu	Leu	260	265	270	
Leu	Tyr	Phe	Trp	Ala	Phe	Thr	Leu	Leu	Cys	Glu	Glu	Leu	Arg	Gln	Gly	275	280	285	
Leu	Ser	Gly	Gly	Gly	Gly	Ser	Leu	Ala	Ser	Gly	Gly	Pro	Gly	Pro	Gly	290	295	300	

-38-

His	Ala	Ser	Leu	Ser	Gln	Arg	Leu	Arg	Leu	Tyr	Leu	Ala	Asp	Ser	Trp
305					310					315					320
Asn	Gln	Cys	Asp	Leu	Val	Ala	Leu	Thr	Cys	Phe	Leu	Leu	Gly	Val	Gly
				325					330					335	
Cys	Arg	Leu	Thr	Pro	Gly	Leu	Tyr	His	Leu	Gly	Arg	Thr	Val	Leu	Cys
				340				345					350		
Ile	Asp	Phe	Met	Val	Phe	Thr	Val	Arg	Leu	Leu	His	Ile	Phe	Thr	Val
	355						360					365			
Asn	Lys	Gln	Leu	Gly	Pro	Lys	Ile	Val	Ile	Val	Ser	Lys	Met	Met	Lys
	370					375					380				
Asp	Val	Phe	Phe	Phe	Leu	Phe	Phe	Leu	Gly	Val	Trp	Leu	Val	Ala	Tyr
385					390					395					400
Gly	Val	Ala	Thr	Glu	Gly	Leu	Leu	Arg	Pro	Arg	Asp	Ser	Asp	Phe	Pro
				405					410					415	
Ser	Ile	Leu	Arg	Arg	Val	Phe	Tyr	Arg	Pro	Tyr	Leu	Gln	Ile	Phe	Gly
				420				425					430		
Gln	Ile	Pro	Gln	Glu	Asp	Met	Asp	Val	Ala	Leu	Met	Glu	His	Ser	Asn
	435						440					445			
Cys	Ser	Ser	Glu	Pro	Gly	Phe	Trp	Ala	His	Pro	Pro	Gly	Ala	Gln	Ala
	450					455					460				
Gly	Thr	Cys	Val	Ser	Gln	Tyr	Ala	Asn	Trp	Leu	Val	Val	Leu	Leu	Leu
465					470					475					480
Val	Ile	Phe	Leu	Leu	Val	Ala	Asn	Ile	Leu	Leu	Val	Asn	Leu	Leu	Ile
				485					490					495	
Ala	Met	Phe	Ser	Tyr	Thr	Phe	Gly	Lys	Val	Gln	Gly	Asn	Ser	Asp	Leu
			500					505					510		
Tyr	Trp	Lys	Ala	Gln	Arg	Tyr	Arg	Leu	Ile	Arg	Glu	Phe	His	Ser	Arg
	515						520					525			
Pro	Ala	Leu	Ala	Pro	Pro	Phe	Ile	Val	Ile	Ser	His	Leu	Arg	Leu	Leu
	530					535					540				
Leu	Arg	Gln	Leu	Cys	Arg	Arg	Pro	Xaa	Ser	Pro	Gln	Pro	Ser	Ser	Pro
545					550				555						560
Ala	Leu	Glu	His	Phe	Arg	Val	Tyr	Leu	Ser	Lys	Glu	Ala	Glu	Arg	Lys
				565					570					575	
Leu	Leu	Thr	Trp	Glu	Ser	Val	His	Lys	Glu	Asn	Phe	Leu	Leu	Ala	Arg
			580					585					590		
Ala	Arg	Asp	Lys	Arg	Glu	Ser	Asp	Ser	Glu	Xaa	Leu	Lys	Arg	Thr	Ser
	595						600					605			
Gln	Lys	Val	Asp	Leu	Ala	Leu	Lys	Gln	Leu	Gly	His	Ile	Arg	Glu	Tyr
	610					615					620				
Glu	Gln	Arg	Leu	Lys	Val	Leu	Glu	Arg	Glu	Val	Gln	Gln	Cys	Thr	Ser
625					630					635					640
Ala	Pro	Ala	Pro	Gly	Gly	Leu	Val	Leu	Glu	Val	Ser	Pro	Met	Ser	Ile
				645					650					655	
Trp	Ala	Thr	Val	Arg	Thr	Thr	Phe	Gly	Ser	Val	Ile	Leu	Thr	Asn	His
			660				665						670		
Ser	Met	Pro	Gly	Ser	Ser	Gln	Asn	Gln	Ser	Gln	Pro	Gly	Arg	Ile	Lys
	675						680					685			
Ala	Trp	Ile	Pro	Gly	Arg	Tyr	Pro	Ser	Gly	Gly	Cys	Arg	Val	Leu	Gly
	690					695					700				
Val	Thr	Gly	Thr	Thr	Asp	Pro	Ser	Pro	Leu	Thr	Asp	Ser	Ser	His	Trp
705					710					715					720
Gly	Asn	Lys	Ala	Ile											
				725											

<210> 27
 <211> 7419
 <212> DNA
 <213> Homo Sapiens

<400> 27

cggggaccga	tccagcctcc	ggactctagc	ctaggctttt	gcaaaaagct	atttaggtga	60
cactatagaa	ggtacgctg	caggtaccgg	tccggaattc	ccgggtcgac	ccacgcgtcc	120
gcagccccgt	cgccggcgga	ggcgggcgcg	ggcgcgtnc	ctgtggccag	tcacccggag	180
gagttggtcg	cacaattatg	aaagactcgg	cttctgctgc	tagcgccgga	gctgagttag	240
ttctgagaa	gtttccctgg	gcgttccttg	tccggcggcc	tctgctgccg	cctccggaga	300
cgcttcccga	tagatggcta	caggccgcgg	aggaggagga	ggtggagttg	ctgcccttcc	360
ggagtccgcc	ccgtgaggag	aatgtcccag	aaatcctgga	tagaaagcac	tttgaccaag	420
agggaaatgt	tatatattat	accaagttcc	aaggaccctc	acagatgcct	tccaggatgt	480
caaatttgct	agcaactcgt	caggtgtttt	tgtggctcgt	tggtaagca	acatgcttgt	540
tttactgcaa	gtcttgccat	gaaatactca	gatgtgaaat	tgggtgacca	ttttaatcag	600
gcaatagaag	aatggtctgt	ggaaaagcat	acagaacaga	gcccacgga	tgcttatgga	660
gtcataaatt	ttcaaggggg	ttctcattcc	tacagagcta	agtatgtgag	gctatcatat	720
gacaccaaac	ctgaagtcat	tctgcaactt	ctgcttaaag	aatggcaaat	ggagttaccc	780
aaacttgcta	tctctgtaca	tgggggcatg	cagaaatttg	agcttcaccc	acgaatcaag	840
cagttgcttg	gaaaaggtct	tattaaagct	gcagttacaa	ctggagcctg	gatttttaact	900
ggaggagtaa	acacaggtgt	ggcaaaaacat	gttgagatg	cctcaaaaga	acatgcttcc	960
agatcatctc	gaaagatttg	cactatcgga	atgctccat	ggggagtgat	tgaaaacaga	1020
aatgatcttg	ttgggagaga	tgtggttgct	ccttatcaaa	ccttattgaa	ccccctgagc	1080
aaattgaatg	ttttgaataa	tctgcattcc	catttcatat	tgggtgatga	tggtcactgtt	1140
ggaaagtatg	gggcggaagt	cagactgaga	agagaacttg	aaaaaactat	taatcagcaa	1200
agaattcatg	ctaggattgg	ccagggtgtc	cctgtggtgg	cacttatatt	tgagggtggg	1260
ccaaatgcta	tcctcacagt	tcttgaatac	cttcaggaaa	gccccctgt	tccagtagtt	1320
gtgtgtgaag	gaacaggcag	agctgcagat	ctgctagcgt	atattcataa	acaaacagaa	1380
gaaggaggga	atcttccctga	tgcagcagag	cccgatatta	tttccactat	caaaaaaaca	1440
tttaactttg	gccagaatga	agcacttcat	ttatttcaaa	cactgatgga	gtgcatgaaa	1500
agaaaggagc	ttatcactgt	tttccatatt	gggtcagatg	aacatcaaga	tatagatgta	1560
gcaatactta	ctgcaactgt	aaaaggtact	aatgcatctg	catttgacca	gcttatcctt	1620
acattggcat	gggtagagat	tgacattgcc	aaaaatcatg	tatttggtta	tggtacagcag	1680
tggctgggtg	gactccttga	acaagctatg	cttgatgctc	ttgtaatgga	tagagtgtga	1740
tttgtaaaac	ttcttattga	aaatggagta	agcatgcata	aattccctac	cattccgaga	1800
ctggaagaac	tttacaacac	taaacaaggt	ccaactaatc	caatgctgtt	tcatcttggt	1860
cgagacgtca	aacagggaaa	tcttcctcca	ggatataaga	tcactctgat	tgatatagga	1920
cttggtattg	aatatctcat	gggaggaacc	tacagatgca	cctatactag	gaaacgtttt	1980
cgattaatat	ataatagtct	tgggtgaaat	aatcggaggt	ctggccgaaa	tacctccagc	2040
agcactcctc	agttgcgaaa	gagtcgatga	tcttttgcca	atagggcaga	taaaaaggaa	2100
aaaatgaggc	ataaccattt	cattaagaca	gcacagccct	tccgaccaaa	gattgatata	2160
gttatggaag	aaggaaagaa	gaaaagaacc	aaagatgaaa	ttgtagacat	tgatgatcca	2220
gaaaccaagc	gctttcctta	tccacttaat	gaacttttaa	tttgggcttg	ccttatgaa	2280
aggcagggtc	tgcccgtttt	tttatggcaa	catggtgaag	aatcaatggc	taaagcata	2340
gttgcttgta	agatctatcg	ttcaatggca	tatgaagcaa	agcagagtga	cctggtagat	2400
gatacttcag	aagaactaaa	acagtattcc	aatgattttg	gtcagttggc	cgttgaaatt	2460
ttagaacagt	ccttcagaca	agatgaaacc	atggctatga	aattgctcac	ttatgaactg	2520
aagaactgga	gtaattcaac	ctgccttaag	ttagcagttt	cttcaagact	tagacctttt	2580
gtagctcaca	cctgtacaca	aatgttgcta	tctgatatgt	ggatgggaag	gctgaatatg	2640
aggaaaaatt	cctgttacaa	ggtcatacta	agcattttag	ttccacctgc	catattgctg	2700
ttagagtata	aaactaaggc	tgaaatgtcc	catatcccac	aatctcaaga	tgctcatcag	2760
atgacaatgg	atgacagcga	aaacaacttt	cagaacataa	cagaagagat	ccccatggaa	2820
gtgtttaaag	aagtacggat	tttggaatg	aatgaaggaa	agaatgagat	ggagatacaa	2880
atgaaatcaa	aaaagcttcc	aattacgcga	aagttttatg	ccttttatca	tgcaccaatt	2940
gtaaaattct	ggtttaacac	gttgccatat	ttaggatttc	tgatgcttta	tacatttgtg	3000
gttcttgtag	aaatggaaac	gttaccttca	gttcaagaat	ggattgttat	tgcttatatt	3060
tttacttatg	ccattgagaa	agtcctgtag	atctttatgt	ctgaagctgg	gaaagttaac	3120
cagaagatta	aagtatgggt	tagtgattac	ttcaacatca	gtgatacaat	tgccataatt	3180
tctttcttca	ttggatttgg	actaagattt	ggagcaaaa	ggaactttgc	aaatgcata	3240
gataatcatg	tttttggtgg	tggaagatta	atttactgtc	ttaacataat	attttggtat	3300
gtgcgtttgc	tagattttct	agctgtaaat	caacaggcag	gaccttatgt	aatgatgatt	3360
ggaaaaatgg	tggccaatat	gttctacatt	gtagtgtatta	tggctcttgt	attacttagt	3420
tttggtgttc	ccagaaaggc	aatactttat	cctcatgaag	caccatcttg	gactcttgct	3480
aaagatatag	tttttcaccc	atactggatg	atttttggtg	aagtttatgc	atacgaaatt	3540

gatgtgtgtg	caaagtattc	tgttatccct	caaattctgt	gtcctgggac	gtgggtgact	3600
ccattttctt	aagcagctta	cctctttgtt	cagtatatca	ttatgggtta	tcttcttatt	3660
gcatttttca	acaatgtgtt	tttacaagt	aaggcaattt	ccaatattgt	atggaagtac	3720
cagcgttatc	attttattat	ggcttatcat	gagaaaccag	ttctgcctcc	tccacttatc	3780
attcttagcc	atatagtttc	tctgttttgc	tgcataatgt	agagaagaaa	gaaagataag	3840
acttccgatg	gaccaaact	tttcttaaca	gaagaagatc	aaaagaaact	tcatgatttt	3900
gaagagcagt	gtgttgaaat	gtatttcaat	gaaaaagatg	acaaatttca	ttctgggagt	3960
gaagagagaa	ttcgtgtcac	ttttgaaaga	gtggaacaga	tgtgcattca	gattaaagaa	4020
gttgagagtc	gtgtcaacta	cataaaaaga	tcattacaat	cattagattc	tcaaattggc	4080
catttgcaag	atctttcagc	cctgacggta	gatacattaa	aaacactcac	tgcccagaaa	4140
gcgtcggaa	ctagcaaat	tcataatgaa	atcacacgag	aactgagcat	ttccaaacac	4200
ttggctcaaa	acctttatga	tgatggctct	gtaagacctt	ctgtatggaa	aaagcatggg	4260
gttgtaata	cacttagctc	ctctcttctt	caagggtgat	ttgaaagtaa	taatcctttt	4320
cattgtaata	ttttaatgaa	agatgacaaa	gatccccagt	gtaatatatt	tggtcaagac	4380
ttacctcgag	taccccgag	aaaagaattt	aattttccag	aggctgggtc	ctcttctggt	4440
gccttattcc	caagtgtgtg	ttccctccca	gaactgcgac	agagactaca	tggggtgaaa	4500
ctcttaaaaa	tatttaataa	aaatcaaaaa	ttaggcagtt	catctactag	cataccacat	4560
ctgtcatccc	caccaacca	attttttgtt	agtacaccat	ctcagccaag	ttgcaaaagc	4620
cacttggaag	ctggaacca	agatcaagaa	actgtttgct	ctaaagctac	agaaggagat	4680
aatacagaat	ttggagcatt	tgtaggacac	agagatagca	tggttttaca	gagggtttaa	4740
gaaacatcaa	acaagataaa	aatactatcc	aataacaata	cttctgaaaa	cactttgaaa	4800
cagtgaggtt	ctcttgctgg	atttactgac	tgctcacagaa	cttccattcc	tggtcattca	4860
aaacaagcag	aaaaaatcag	tagaaggcca	tcacccgaag	acactcatga	agtagattcc	4920
aaagcagctt	taataccgga	ttgggttaca	gatagaccat	caaacagaga	aatgccatct	4980
gaagaaggaa	cattaaatgg	tctcacttct	ccattttaagc	cagctatgga	tacaatttac	5040
tattattcag	ctgtggaaa	aaataacttg	atgaggttat	cacagagcat	tccattttaca	5100
cctgtgcctc	caagagggga	gcctgtcaca	gtgtatcggt	tggaagagag	ttcacccaac	5160
atactaaata	acagcatgtc	ttcttggtca	caactaggcc	tctgtgccaa	aatagagttt	5220
ttaaagcaag	aggagatggg	aggaggttta	cgaagagctg	tcaaagtaca	gtgtacctgg	5280
tcagaacatg	atatcctcaa	atcagggcac	ctttatatta	tcaaattctt	tcttccagag	5340
gtggtaata	catggtcaag	tattttataa	gaagatacac	ttctgcatct	ctgtctgaga	5400
gaaattcaac	aacagagagc	agcacaaaag	cttacggttg	cctttaatca	aatgaaaccc	5460
aaatccatac	catattctcc	aagggttcctt	gaagttttcc	tgtgtatttg	ccattcagca	5520
ggacagtggg	ttgctgtgga	agaatgtatg	actggagaat	ttagaaaata	caacaataat	5580
aatggagatg	agattattcc	aactaatact	ctggaagaga	tcatgctagc	ctttagccac	5640
tggaacttac	aatatacaag	aggggagtta	ctgggtactg	atttgcaagg	tggtgggtgaa	5700
aatttgactg	acccatctgt	gataaaaagca	gaagaaaaga	gatcctgtga	tatgggtttt	5760
ggcccagcaa	atctaggaga	agatgcaatt	aaaaaacttca	gagcaaaaaca	tcactgtaat	5820
tcttgctgta	gaaagcttaa	acttccagat	ctggaagaga	atgattatac	gcctgataaa	5880
attatatttc	ctcaggatga	gccttcagat	ttgaatcttc	agcctggaaa	ttccaccaaa	5940
gaatcagaat	caactaatc	tggtcgtctg	atgtttataat	attaatatta	ctgaatcatt	6000
ggttttgcct	gcacctcaca	gaaatgttac	tgtgtcactt	ttccctcggg	aggaaattgt	6060
ttggtaatat	agaaaggtgt	atgcaagtgt	aatgtgtctg	ctccagcaca	gttaaaagg	6120
caatattctt	ttgacctgat	taatcagtca	gaaagtccct	ataggataga	gctggcagct	6180
gagaaatttt	aaaggttaatt	gataattagt	atgtgttaact	ttttaaagg	ctctttgtat	6240
agcagaggat	ctcatttgac	tttgttttga	tgaggggtgat	gccctctctt	atgtgggtaca	6300
ataccattaa	ccaaaggtag	gtgtccatgc	agatttttatt	ggcagctggt	ttattgccat	6360
tcaactaggg	aaatgaagaa	atcacgcagc	cttttggtta	aatggcagtc	aaaattttcc	6420
tcagtgtatt	tagtgtgttc	agtgatgata	tcactgggtc	ccaactagat	gcttggtggc	6480
cacgggaagg	gaaatgactt	gttctaattc	taggttcaca	gaggtatgag	aagcctgaac	6540
tgaagaccat	tttcaagagg	gacggtattt	atgaatcagg	gttaggctcc	atattttaa	6600
atagagccag	ttttttttt	aaatagaacc	caaattgtgt	aaaaatgtta	attgggtttt	6660
ttaaacattg	ttttatcaag	tcactgttaa	gtagaagaaa	gccatggtta	actgatacat	6720
aacctaaatt	ataaaagcag	aaacctaact	cactcgtcaa	gggaagttac	cttttgagga	6780
aagttaaaagt	acttttttcc	ctatctgtat	ctatagcaac	aaccagaaac	ttacaaactt	6840
ctccaaagat	tttatttgatt	gttatatcaa	atcagaatgt	aaacatgaac	tcttgcatat	6900
atttaaaatt	gtgttggaac	atttgaacat	gaatgctgtt	tggtgactta	agaaattrat	6960
tcagtnngat	tatcattatg	tganactggc	agattgcagt	gcanccttat	gccaataaaa	7020
tgtaatttar	cagcccagaa	tattgttgaa	tattcaacaa	taacaagaaa	agcttttcat	7080
ctaagtttta	tgctttaatt	tttttctttt	tttttctttt	ttcttttgtt	tccttggtac	7140

-41-

taattttaat	ttttatttgg	aagggagcag	tataaagctt	atttgtattt	agtagtgtat	7200
ctcatagata	cagacaaggc	aagagatgat	aagctgttta	aatagtgktt	aatattgatt	7260
gggggtgggg	agaaagaaaa	agtgtattac	ttaaagatac	tatatacskt	ttktatatca	7320
ttaaatcttt	aaaagaaatn	naataaattt	attgttttnc	aaaaaaaaac	ccnntaaaaa	7380
aaaaagggcg	gcccctctag	aggatccctc	gagggggccc			7419

<210> 28
 <211> 1865
 <212> PRT
 <213> Homo Sapiens

<400> 28

Met	Ser	Gln	Lys	Ser	Trp	Ile	Glu	Ser	Thr	Leu	Thr	Lys	Arg	Glu	Cys
1				5					10					15	
Val	Tyr	Ile	Ile	Pro	Ser	Ser	Lys	Asp	Pro	His	Arg	Cys	Leu	Pro	Gly
			20					25					30		
Cys	Gln	Ile	Cys	Gln	Gln	Leu	Val	Arg	Cys	Phe	Cys	Gly	Arg	Leu	Val
		35					40					45			
Lys	Gln	His	Ala	Cys	Phe	Thr	Ala	Ser	Leu	Ala	Met	Lys	Tyr	Ser	Asp
		50				55					60				
Val	Lys	Leu	Gly	Asp	His	Phe	Asn	Gln	Ala	Ile	Glu	Glu	Trp	Ser	Val
					70					75					80
Glu	Lys	His	Thr	Glu	Gln	Ser	Pro	Thr	Asp	Ala	Tyr	Gly	Val	Ile	Asn
				85					90					95	
Phe	Gln	Gly	Gly	Ser	His	Ser	Tyr	Arg	Ala	Lys	Tyr	Val	Arg	Leu	Ser
			100					105					110		
Tyr	Asp	Thr	Lys	Pro	Glu	Val	Ile	Leu	Gln	Leu	Leu	Leu	Lys	Glu	Trp
			115				120					125			
Gln	Met	Glu	Leu	Pro	Lys	Leu	Val	Ile	Ser	Val	His	Gly	Gly	Met	Gln
			130			135					140				
Lys	Phe	Glu	Leu	His	Pro	Arg	Ile	Lys	Gln	Leu	Leu	Gly	Lys	Gly	Leu
					150					155					160
Ile	Lys	Ala	Ala	Val	Thr	Gly	Ala	Trp	Ile	Leu	Thr	Gly	Gly	Val	
				165				170					175		
Asn	Thr	Gly	Val	Ala	Lys	His	Val	Gly	Asp	Ala	Leu	Lys	Glu	His	Ala
			180					185					190		
Ser	Arg	Ser	Ser	Arg	Lys	Ile	Cys	Thr	Ile	Gly	Ile	Ala	Pro	Trp	Gly
			195				200					205			
Val	Ile	Glu	Asn	Arg	Asn	Asp	Leu	Val	Gly	Arg	Asp	Val	Val	Ala	Pro
			210			215					220				
Tyr	Gln	Thr	Leu	Leu	Asn	Pro	Leu	Ser	Lys	Leu	Asn	Val	Leu	Asn	Asn
					230					235					240
Leu	His	Ser	His	Phe	Ile	Leu	Val	Asp	Asp	Gly	Thr	Val	Gly	Lys	Tyr
				245					250					255	
Gly	Ala	Glu	Val	Arg	Leu	Arg	Arg	Glu	Leu	Glu	Lys	Thr	Ile	Asn	Gln
			260					265					270		
Gln	Arg	Ile	His	Ala	Arg	Ile	Gly	Gln	Gly	Val	Pro	Val	Val	Ala	Leu
							280					285			
Ile	Phe	Glu	Gly	Gly	Pro	Asn	Val	Ile	Leu	Thr	Val	Leu	Glu	Tyr	Leu
						295					300				
Gln	Glu	Ser	Pro	Pro	Val	Pro	Val	Val	Val	Cys	Glu	Gly	Thr	Gly	Arg
					310					315					320
Ala	Ala	Asp	Leu	Leu	Ala	Tyr	Ile	His	Lys	Gln	Thr	Glu	Glu	Gly	Gly
				325					330					335	
Asn	Leu	Pro	Asp	Ala	Ala	Glu	Pro	Asp	Ile	Ile	Ser	Thr	Ile	Lys	Lys
			340					345					350		
Thr	Phe	Asn	Phe	Gly	Gln	Asn	Glu	Ala	Leu	His	Leu	Phe	Gln	Thr	Leu
			355				360					365			

-42-

Met	Glu	Cys	Met	Lys	Arg	Lys	Glu	Leu	Ile	Thr	Val	Phe	His	Ile	Gly
370						375					380				
Ser	Asp	Glu	His	Gln	Asp	Ile	Asp	Val	Ala	Ile	Leu	Thr	Ala	Leu	Leu
385					390					395					400
Lys	Gly	Thr	Asn	Ala	Ser	Ala	Phe	Asp	Gln	Leu	Ile	Leu	Thr	Leu	Ala
			405						410						415
Trp	Asp	Arg	Val	Asp	Ile	Ala	Lys	Asn	His	Val	Phe	Val	Tyr	Gly	Gln
			420					425					430		
Gln	Trp	Leu	Val	Gly	Ser	Leu	Glu	Gln	Ala	Met	Leu	Asp	Ala	Leu	Val
		435					440					445			
Met	Asp	Arg	Val	Ala	Phe	Val	Lys	Leu	Leu	Ile	Glu	Asn	Gly	Val	Ser
	450					455					460				
Met	His	Lys	Phe	Leu	Thr	Ile	Pro	Arg	Leu	Glu	Glu	Leu	Tyr	Asn	Thr
465					470					475					480
Lys	Gln	Gly	Pro	Thr	Asn	Pro	Met	Leu	Phe	His	Leu	Val	Arg	Asp	Val
				485					490					495	
Lys	Gln	Gly	Asn	Leu	Pro	Pro	Gly	Tyr	Lys	Ile	Thr	Leu	Ile	Asp	Ile
			500					505					510		
Gly	Leu	Val	Ile	Glu	Tyr	Leu	Met	Gly	Gly	Thr	Tyr	Arg	Cys	Thr	Tyr
	515						520					525			
Thr	Arg	Lys	Arg	Phe	Arg	Leu	Ile	Tyr	Asn	Ser	Leu	Gly	Gly	Asn	Asn
	530					535					540				
Arg	Arg	Ser	Gly	Arg	Asn	Thr	Ser	Ser	Ser	Thr	Pro	Gln	Leu	Arg	Lys
545					550					555					560
Ser	His	Glu	Ser	Phe	Gly	Asn	Arg	Ala	Asp	Lys	Lys	Glu	Lys	Met	Arg
			565						570					575	
His	Asn	His	Phe	Ile	Lys	Thr	Ala	Gln	Pro	Phe	Arg	Pro	Lys	Ile	Asp
			580					585					590		
Thr	Val	Met	Glu	Glu	Gly	Lys	Lys	Arg	Thr	Lys	Asp	Glu	Ile	Val	
	595						600				605				
Asp	Ile	Asp	Asp	Pro	Glu	Thr	Lys	Arg	Phe	Pro	Tyr	Pro	Leu	Asn	Glu
	610					615					620				
Leu	Leu	Ile	Trp	Ala	Cys	Leu	Met	Lys	Arg	Gln	Val	Met	Ala	Arg	Phe
625					630					635					640
Leu	Trp	Gln	His	Gly	Glu	Glu	Ser	Met	Ala	Lys	Ala	Leu	Val	Ala	Cys
				645					650					655	
Lys	Ile	Tyr	Arg	Ser	Met	Ala	Tyr	Glu	Ala	Lys	Gln	Ser	Asp	Leu	Val
			660					665					670		
Asp	Asp	Thr	Ser	Glu	Glu	Leu	Lys	Gln	Tyr	Ser	Asn	Asp	Phe	Gly	Gln
		675					680					685			
Leu	Ala	Val	Glu	Leu	Leu	Glu	Gln	Ser	Phe	Arg	Gln	Asp	Glu	Thr	Met
	690					695					700				
Ala	Met	Lys	Leu	Leu	Thr	Tyr	Glu	Leu	Lys	Asn	Trp	Ser	Asn	Ser	Thr
705					710					715					720
Cys	Leu	Lys	Leu	Ala	Val	Ser	Ser	Arg	Leu	Arg	Pro	Phe	Val	Ala	His
				725					730					735	
Thr	Cys	Thr	Gln	Met	Leu	Leu	Ser	Asp	Met	Trp	Met	Gly	Arg	Leu	Asn
			740					745				750			
Met	Arg	Lys	Asn	Ser	Trp	Tyr	Lys	Val	Ile	Leu	Ser	Ile	Leu	Val	Pro
		755					760					765			
Pro	Ala	Ile	Leu	Leu	Leu	Glu	Tyr	Lys	Thr	Lys	Ala	Glu	Met	Ser	His
		770				775					780				
Ile	Pro	Gln	Ser	Gln	Asp	Ala	His	Gln	Met	Thr	Met	Asp	Asp	Ser	Glu
785					790					795					800
Asn	Asn	Phe	Gln	Asn	Ile	Thr	Glu	Glu	Ile	Pro	Met	Glu	Val	Phe	Lys
				805					810					815	
Glu	Val	Arg	Ile	Leu	Asp	Ser	Asn	Glu	Gly	Lys	Asn	Glu	Met	Glu	Ile
			820					825				830			
Gln	Met	Lys	Ser	Lys	Lys	Leu	Pro	Ile	Thr	Arg	Lys	Phe	Tyr	Ala	Phe
		835					840					845			

-43-

Tyr	His	Ala	Pro	Ile	Val	Lys	Phe	Trp	Phe	Asn	Thr	Leu	Ala	Tyr	Leu
	850					855				860					
Gly	Phe	Leu	Met	Leu	Tyr	Thr	Phe	Val	Val	Leu	Val	Gln	Met	Glu	Gln
865					870					875					880
Leu	Pro	Ser	Val	Gln	Glu	Trp	Ile	Val	Ile	Ala	Tyr	Ile	Phe	Thr	Tyr
				885						890					895
Ala	Ile	Glu	Lys	Val	Arg	Glu	Ile	Phe	Met	Ser	Glu	Ala	Gly	Lys	Val
			900						905					910	
Asn	Gln	Lys	Ile	Lys	Val	Trp	Phe	Ser	Asp	Tyr	Phe	Asn	Ile	Ser	Asp
		915					920					925			
Thr	Ile	Ala	Ile	Ile	Ser	Phe	Phe	Ile	Gly	Phe	Gly	Leu	Arg	Phe	Gly
930						935					940				
Ala	Lys	Trp	Asn	Phe	Ala	Asn	Ala	Tyr	Asp	Asn	His	Val	Phe	Val	Ala
945					950					955					960
Gly	Arg	Leu	Ile	Tyr	Cys	Leu	Asn	Ile	Ile	Phe	Trp	Tyr	Val	Arg	Leu
				965						970					975
Leu	Asp	Phe	Leu	Ala	Val	Asn	Gln	Gln	Ala	Gly	Pro	Tyr	Val	Met	Met
			980						985					990	
Ile	Gly	Lys	Met	Val	Ala	Asn	Met	Phe	Tyr	Ile	Val	Val	Ile	Met	Ala
			995				1000						1005		
Leu	Val	Leu	Leu	Ser	Phe	Gly	Val	Pro	Arg	Lys	Ala	Ile	Leu	Tyr	Pro
							1015					1020			
His	Glu	Ala	Pro	Ser	Trp	Thr	Leu	Ala	Lys	Asp	Ile	Val	Phe	His	Pro
1025					1030					1035					1040
Tyr	Trp	Met	Ile	Phe	Gly	Glu	Val	Tyr	Ala	Tyr	Glu	Ile	Asp	Val	Cys
				1045						1050					1055
Ala	Asn	Asp	Ser	Val	Ile	Pro	Gln	Ile	Cys	Gly	Pro	Gly	Thr	Trp	Leu
				1060					1065						1070
Thr	Pro	Phe	Leu	Gln	Ala	Val	Tyr	Leu	Phe	Val	Gln	Tyr	Ile	Ile	Met
			1075					1080					1085		
Val	Asn	Leu	Leu	Ile	Ala	Phe	Phe	Asn	Asn	Val	Tyr	Leu	Gln	Val	Lys
			1090			1095					1100				
Ala	Ile	Ser	Asn	Ile	Val	Trp	Lys	Tyr	Gln	Arg	Tyr	His	Phe	Ile	Met
1105					1110					1115					1120
Ala	Tyr	His	Glu	Lys	Pro	Val	Leu	Pro	Pro	Leu	Ile	Ile	Leu	Ser	
				1125						1130				1135	
His	Ile	Val	Ser	Leu	Phe	Cys	Cys	Ile	Cys	Lys	Arg	Arg	Lys	Lys	Asp
			1140						1145					1150	
Lys	Thr	Ser	Asp	Gly	Pro	Lys	Leu	Phe	Leu	Thr	Glu	Glu	Asp	Gln	Lys
			1155				1160						1165		
Lys	Leu	His	Asp	Phe	Glu	Glu	Gln	Cys	Val	Glu	Met	Tyr	Phe	Asn	Glu
			1170				1175				1180				
Lys	Asp	Asp	Lys	Phe	His	Ser	Gly	Ser	Glu	Glu	Arg	Ile	Arg	Val	Thr
1185					1190					1195					1200
Phe	Glu	Arg	Val	Glu	Gln	Met	Cys	Ile	Gln	Ile	Lys	Glu	Val	Gly	Asp
				1205						1210					1215
Arg	Val	Asn	Tyr	Ile	Lys	Arg	Ser	Leu	Gln	Ser	Leu	Asp	Ser	Gln	Ile
				1220					1225					1230	
Gly	His	Leu	Gln	Asp	Leu	Ser	Ala	Leu	Thr	Val	Asp	Thr	Leu	Lys	Thr
				1235				1240						1245	
Leu	Thr	Ala	Gln	Lys	Ala	Ser	Glu	Ala	Ser	Lys	Val	His	Asn	Glu	Ile
				1250			1255					1260			
Thr	Arg	Glu	Leu	Ser	Ile	Ser	Lys	His	Leu	Ala	Gln	Asn	Leu	Ile	Asp
1265					1270					1275					1280
Asp	Gly	Pro	Val	Arg	Pro	Ser	Val	Trp	Lys	Lys	His	Gly	Val	Val	Asn
				1285						1290					1295
Thr	Leu	Ser	Ser	Leu	Pro	Gln	Gly	Asp	Leu	Glu	Ser	Asn	Asn	Pro	
				1300				1305					1310		
Phe	His	Cys	Asn	Ile	Leu	Met	Lys	Asp	Asp	Lys	Asp	Pro	Gln	Cys	Asn
			1315					1320							1325

-44-

Ile Phe Gly Gln Asp Leu Pro Ala Val Pro Gln Arg Lys Glu Phe Asn
 1330 1335 1340
 Phe Pro Glu Ala Gly Ser Ser Ser Gly Ala Leu Phe Pro Ser Ala Val
 1345 1350 1355 1360
 Ser Pro Pro Glu Leu Arg Gln Arg Leu His Gly Val Glu Leu Leu Lys
 1365 1370 1375
 Ile Phe Asn Lys Asn Gln Lys Leu Gly Ser Ser Ser Thr Ser Ile Pro
 1380 1385 1390
 His Leu Ser Ser Pro Pro Thr Lys Phe Phe Val Ser Thr Pro Ser Gln
 1395 1400 1405
 Pro Ser Cys Lys Ser His Leu Glu Thr Gly Thr Lys Asp Gln Glu Thr
 1410 1415 1420
 Val Cys Ser Lys Ala Thr Glu Gly Asp Asn Thr Glu Phe Gly Ala Phe
 1425 1430 1435 1440
 Val Gly His Arg Asp Ser Met Asp Leu Gln Arg Phe Lys Glu Thr Ser
 1445 1450 1455
 Asn Lys Ile Lys Ile Leu Ser Asn Asn Asn Thr Ser Glu Asn Thr Leu
 1460 1465 1470
 Lys Arg Val Ser Ser Leu Ala Gly Phe Thr Asp Cys His Arg Thr Ser
 1475 1480 1485
 Ile Pro Val His Ser Lys Gln Ala Glu Lys Ile Ser Arg Arg Pro Ser
 1490 1495 1500
 Thr Glu Asp Thr His Glu Val Asp Ser Lys Ala Ala Leu Ile Pro Asp
 1505 1510 1515 1520
 Trp Leu Gln Asp Arg Pro Ser Asn Arg Glu Met Pro Ser Glu Glu Gly
 1525 1530 1535
 Thr Leu Asn Gly Leu Thr Ser Pro Phe Lys Pro Ala Met Asp Thr Asn
 1540 1545 1550
 Tyr Tyr Tyr Ser Ala Val Glu Arg Asn Asn Leu Met Arg Leu Ser Gln
 1555 1560 1565
 Ser Ile Pro Phe Thr Pro Val Pro Pro Arg Gly Glu Pro Val Thr Val
 1570 1575 1580
 Tyr Arg Leu Glu Glu Ser Ser Pro Asn Ile Leu Asn Asn Ser Met Ser
 1585 1590 1595 1600
 Ser Trp Ser Gln Leu Gly Leu Cys Ala Lys Ile Glu Phe Leu Ser Lys
 1605 1610 1615
 Glu Glu Met Gly Gly Gly Leu Arg Arg Ala Val Lys Val Gln Cys Thr
 1620 1625 1630
 Trp Ser Glu His Asp Ile Leu Lys Ser Gly His Leu Tyr Ile Ile Lys
 1635 1640 1645
 Ser Phe Leu Pro Glu Val Val Asn Thr Trp Ser Ser Ile Tyr Lys Glu
 1650 1655 1660
 Asp Thr Val Leu His Leu Cys Leu Arg Glu Ile Gln Gln Gln Arg Ala
 1665 1670 1675 1680
 Ala Gln Lys Leu Thr Phe Ala Phe Asn Gln Met Lys Pro Lys Ser Ile
 1685 1690 1695
 Pro Tyr Ser Pro Arg Phe Leu Glu Val Phe Leu Leu Tyr Cys His Ser
 1700 1705 1710
 Ala Gly Gln Trp Phe Ala Val Glu Glu Cys Met Thr Gly Glu Phe Arg
 1715 1720 1725
 Lys Tyr Asn Asn Asn Asn Gly Asp Glu Ile Ile Pro Thr Asn Thr Leu
 1730 1735 1740
 Glu Glu Ile Met Leu Ala Phe Ser His Trp Thr Tyr Glu Tyr Thr Arg
 1745 1750 1755 1760
 Gly Glu Leu Leu Val Leu Asp Leu Gln Gly Val Gly Glu Asn Leu Thr
 1765 1770 1775
 Asp Pro Ser Val Ile Lys Ala Glu Glu Lys Arg Ser Cys Asp Met Val
 1780 1785 1790
 Phe Gly Pro Ala Asn Leu Gly Glu Asp Ala Ile Lys Asn Phe Arg Ala
 1795 1800 1805

-45-

Lys His His Cys Asn Ser Cys Cys Arg Lys Leu Lys Leu Pro Asp Leu
 1810 1815 1820
 Lys Arg Asn Asp Tyr Thr Pro Asp Lys Ile Ile Phe Pro Gln Asp Glu
 1825 1830 1835 1840
 Pro Ser Asp Leu Asn Leu Gln Pro Gly Asn Ser Thr Lys Glu Ser Glu
 1845 1850 1855
 Ser Thr Asn Ser Val Arg Leu Met Leu
 1860 1865

<210> 29
 <211> 4061
 <212> DNA
 <213> Homo Sapiens

<400> 29
 ggtctggaag cagagccggc ggaggggagcg ccggggccct gggctgcagg aggttgccgc 60
 ggccgcggca gcatggtggt gccggagaag gagcagagct ggatcccca gatcttcaag 120
 aagaagacct gcacgacgtt catagttgac tccacagatc cgggagggac cttgtgccag 180
 tgtggggcgcc ccgggaccgc ccaccccgca gtggccatgg aggatgcctt cggggcagcc 240
 gtggtgaccg tgtgggacag cgatgcacac accacggaga agcccaccga tgcctacgga 300
 gagctggact tcacgggggc cgcccgcaag cacagcaatt tctccggct ctctgaccga 360
 acgatccag ctgcagtta tagtctggtc acacgcacat ggggcttccg tccccgaac 420
 ctggtggtgt cagtgtctgg gggatcgggg ggccccgtcc tccagacctg gctgcaggac 480
 ctgctgcgtc gtgggctggt gcgggctgcc cagagcacag gagcctggat tgtcactggg 540
 ggtctgcaca cgggcatcgg ccggcatggt ggtgtggctg tacgggacca tcagatggcc 600
 agcactgggg gcaccaaggt ggtggccatg ggtgtggccc cctggggtgt ggtccggaat 660
 agagacaccc tcatcaaccc caagggctcg ttccctgcga ggtaccggtg gcgcggtgac 720
 ccggaggacg ggggtccagt tcccctggac tacaactact cggccttctt cctggtggac 780
 gacggcacac acggctgcct ggggggcgag aaccgcttcc gtttgcgcct ggagtcctac 840
 atctcacagc agaagacggg cgtgggaggg actggaattg acatccctgt cctgtcctc 900
 ctgattgatg gtgatgagaa gatgttgacg cgaatagaga acgccacca ggctcagctc 960
 ccatgtctcc tcgtggctgg ctacggggga gctgcggact gcctggcgga gaccctggaa 1020
 gacactctgg cccagggag tgggggagcc aggcaaggcg aagcccgaga tcgaatcagg 1080
 cgtttcttcc ccaagggga ccttgaggtc ctgcaggccc aggtggagag gattatgacc 1140
 cggaaggagc tcctgacagt ctattcttct gaggatgggt ctgaggaatt cgagaccata 1200
 gttttgaagg cccttgtgaa ggcctgtggg agctcggagg cctcagccta cctggatgag 1260
 ctgcgtttgg ctgtggcttg gaaccgcgtg gacattgccc agagtgaact ctttcggggg 1320
 gacatccaat ggcggtcctt ccattctgaa gcttccctca tggacgccct gctgaatgac 1380
 cggcctgagt tcgtgcgtt gctcatttcc caccgctca gcctgggcca cttectgacc 1440
 ccgatgcgcc tggcccaact ctacagcgcg gcgcctcca actcgtcat ccgcaacctt 1500
 ttggaccagg cgtccacag cgcaggcacc aaagccccag ccctaaaagg gggagctgcg 1560
 gagctccggc cccctgacgt ggggcatgtg ctgaggatgc tgctggggaa gatgtgcgcg 1620
 ccgaggtacc cctccggggg cgcctgggac cctcaccag gccagggtt cggggagagc 1680
 atgtatctgc tctcggaaca ggccacctcg ccgctctcgc tggatgctgg cctcgggcag 1740
 gccccctgga gcgacctgct tctttgggca ctgttgctga acagggcaca gatggccatg 1800
 tacttctggg agatgggttc caatgcagtt tctcagctc ttggggcctg tttgctgctc 1860
 cgggtgatgg cagcctgga gcctgacgt gagggagcag cacggaggaa agacctggcg 1920
 ttcaagtttg aggggatggg cgttgacctc tttggcgagt gctatcgag cagtgggtg 1980
 agggctgccc gcctcctcct ccgtcgtgc ccgctctggg gggatgccac ttgcctccag 2040
 ctggccatgc aagctgacgc ccgtgccttc tttggccagg atggggtaca gtctctgctc 2100
 acacagaagt ggtggggaga tatggccagc actacacca tctgggccct gggtctcgcc 2160
 ttcttttgcc ctccactcat ctacaccgc ctcatcacct tcaggaaatc agaagaggag 2220
 ccacacggg aggagctaga gtttgacatg gatagtgtca ttaatgggga agggcctgtc 2280
 gggacggcg acccagccga gaagacgccg ctgggggtcc cgcgccagtc gggccgtccg 2340
 ggttgctgcg ggggcccgtg cggggggcg cgggtgcctac gccgctggtt ccacttctgg 2400
 ggcgcgcgg tgacctctt catgggcaac gtggctgact acctgctgtt cctgctctt 2460
 ttctcgcggt gctgctcgt ggatttccag ccggcgccg ccggctcctt ggagctgctg 2520
 ctctatttct gggctttcac gctgctgtgc gaggaactgc gccagggcct gagcggaggc 2580
 gggggcagcc tcgccagcg gggccccgg cctggccatg cctcactgag ccagcgctg 2640
 cgcctctacc tcgccgacag ctggaaccag tgcgacctag tggctctcac ctgcttctc 2700

-46-

```

ctgggcggtg gctgccggct gaccccggtt ttgtaccacc tgggcccgcac tgtcctctgc 2760
atcgacttca tggttttcac ggtgcggtcg cttcacatct tcacgggtcaa caaacagctg 2820
gggccaaga tcgtcatcgt gagcaagatg atgaaggacg tgttcttctt cctcttcttc 2880
ctcggcgtgt ggctggtagc ctatggcggt gccacggagg ggctcctgag gccacgggac 2940
agtgaattcc caagtatcct gcgccgcgtc ttctaccgtc cctacctgca gatcttcggg 3000
cagattcccc aggaggacat ggacgtggcc ctcatggagc acagcaactg ctcgtcggag 3060
cccggcttct gggcacaccc tcctggggcc caggcgggca cctgcgtctc ccagtatgcc 3120
aactggctgg tgggtgctgt cctcgtcacg ttctgctcgt tggccaacat cctgctggtc 3180
aacttgctca ttgccatgtt cagttacaca ttccggcaaag tacagggcaa cagcgatctc 3240
tactggaagg cgcagcgtaa ccgcctcatc cgggaattcc actctcggcc cgcgctggcc 3300
ccgcccctta tcgtcatctc ccacttgccg ctcctgctca ggcaattgtg caggcgaccc 3360
cggagccccc agccgtcctc cccggccctc gagcatttcc gggtttacct ttctaaggaa 3420
gccgagcgga agctgctaac gtgggaatcg gtgcataagg agaactttct gctggcacgc 3480
gctagggaca agcgggagag cgactccgag cgtctgaagc gcacgtccca gaaggtggac 3540
ttggcactga aacagctggg acacatccgc gagtacgaac agcgcctgaa agtgctggag 3600
cgggaggtcc agcagtgtag ccgcgtcctg ggggtgggtg ccgaggccct gagccgctc 3660
gccttgctgc ccccaggtg gcgcgccacc cctgacctgc ctgggtccaa agactgagcc 3720
ctgctggcgg acttcaagga gaagccccc caggggattt tgctcctaga gtaaggctca 3780
tctgggcctc ggcccccgca cctggtggcc ttgtccttga ggtgagcccc atgtccatct 3840
gggccactgt caggaccacc tttgggagtg tcacacctac aaaccacagc atgcccggct 3900
cctccagaaa ccagtcacg cctgggagga tcaaggcctg gatcccgggc cgttatccat 3960
ctggaggtcg cagggtcctt ggggtaacag ggaccacaga cccctcacca ctcacagatt 4020
cctcacactg gggaaataaa gccatttcag agggaaaaaa a 4061

```

<210> 30
 <211> 1214
 <212> PRT
 <213> Homo Sapiens

<400> 30

Met	Val	Val	Pro	Glu	Lys	Glu	Gln	Ser	Trp	Ile	Pro	Lys	Ile	Phe	Lys
1				5				10						15	
Lys	Lys	Thr	Cys	Thr	Thr	Phe	Ile	Val	Asp	Ser	Thr	Asp	Pro	Gly	Gly
			20					25					30		
Thr	Leu	Cys	Gln	Cys	Gly	Arg	Pro	Arg	Thr	Ala	His	Pro	Ala	Val	Ala
			35				40					45			
Met	Glu	Asp	Ala	Phe	Gly	Ala	Ala	Val	Val	Thr	Val	Trp	Asp	Ser	Asp
	50				55						60				
Ala	His	Thr	Thr	Glu	Lys	Pro	Thr	Asp	Ala	Tyr	Gly	Glu	Leu	Asp	Phe
	65			70				75						80	
Thr	Gly	Ala	Gly	Arg	Lys	His	Ser	Asn	Phe	Leu	Arg	Leu	Ser	Asp	Arg
			85					90						95	
Thr	Asp	Pro	Ala	Ala	Val	Tyr	Ser	Leu	Val	Thr	Arg	Thr	Trp	Gly	Phe
			100					105					110		
Arg	Ala	Pro	Asn	Leu	Val	Val	Ser	Val	Leu	Gly	Gly	Ser	Gly	Gly	Pro
			115				120					125			
Val	Leu	Gln	Thr	Trp	Leu	Gln	Asp	Leu	Leu	Arg	Arg	Gly	Leu	Val	Arg
			130			135					140				
Ala	Ala	Gln	Ser	Thr	Gly	Ala	Trp	Ile	Val	Thr	Gly	Gly	Leu	His	Thr
					150					155				160	
Gly	Ile	Gly	Arg	His	Val	Gly	Val	Ala	Val	Arg	Asp	His	Gln	Met	Ala
				165				170						175	
Ser	Thr	Gly	Gly	Thr	Lys	Val	Val	Ala	Met	Gly	Val	Ala	Pro	Trp	Gly
			180					185					190		
Val	Val	Arg	Asn	Arg	Asp	Thr	Leu	Ile	Asn	Pro	Lys	Gly	Ser	Phe	Pro
			195				200					205			
Ala	Arg	Tyr	Arg	Trp	Arg	Gly	Asp	Pro	Glu	Asp	Gly	Val	Gln	Phe	Pro
			210			215					220				
Leu	Asp	Tyr	Asn	Tyr	Ser	Ala	Phe	Phe	Leu	Val	Asp	Asp	Gly	Thr	His
			225			230				235					240

-47-

Gly	Cys	Leu	Gly	Gly	Glu	Asn	Arg	Phe	Arg	Leu	Arg	Leu	Glu	Ser	Tyr
				245					250					255	
Ile	Ser	Gln	Gln	Lys	Thr	Gly	Val	Gly	Gly	Thr	Gly	Ile	Asp	Ile	Pro
			260					265					270		
Val	Leu	Leu	Leu	Leu	Ile	Asp	Gly	Asp	Glu	Lys	Met	Leu	Thr	Arg	Ile
			275				280					285			
Glu	Asn	Ala	Thr	Gln	Ala	Gln	Leu	Pro	Cys	Leu	Leu	Val	Ala	Gly	Ser
	290					295				300					
Gly	Gly	Ala	Ala	Asp	Cys	Leu	Ala	Glu	Thr	Leu	Glu	Asp	Thr	Leu	Ala
305				310						315					320
Pro	Gly	Ser	Gly	Gly	Ala	Arg	Gln	Gly	Glu	Ala	Arg	Asp	Arg	Ile	Arg
			325					330						335	
Arg	Phe	Phe	Pro	Lys	Gly	Asp	Leu	Glu	Val	Leu	Gln	Ala	Gln	Val	Glu
			340				345						350		
Arg	Ile	Met	Thr	Arg	Lys	Glu	Leu	Leu	Thr	Val	Tyr	Ser	Ser	Glu	Asp
		355				360						365			
Gly	Ser	Glu	Glu	Phe	Glu	Thr	Ile	Val	Leu	Lys	Ala	Leu	Val	Lys	Ala
	370					375				380					
Cys	Gly	Ser	Ser	Glu	Ala	Ser	Ala	Tyr	Leu	Asp	Glu	Leu	Arg	Leu	Ala
385				390					395						400
Val	Ala	Trp	Asn	Arg	Val	Asp	Ile	Ala	Gln	Ser	Glu	Leu	Phe	Arg	Gly
			405					410						415	
Asp	Ile	Gln	Trp	Arg	Ser	Phe	His	Leu	Glu	Ala	Ser	Leu	Met	Asp	Ala
		420					425						430		
Leu	Leu	Asn	Asp	Arg	Pro	Glu	Phe	Val	Arg	Leu	Leu	Ile	Ser	His	Gly
		435				440						445			
Leu	Ser	Leu	Gly	His	Phe	Leu	Thr	Pro	Met	Arg	Leu	Ala	Gln	Leu	Tyr
	450				455					460					
Ser	Ala	Ala	Pro	Ser	Asn	Ser	Leu	Ile	Arg	Asn	Leu	Leu	Asp	Gln	Ala
465				470					475						480
Ser	His	Ser	Ala	Gly	Thr	Lys	Ala	Pro	Ala	Leu	Lys	Gly	Gly	Ala	Ala
			485					490						495	
Glu	Leu	Arg	Pro	Pro	Asp	Val	Gly	His	Val	Leu	Arg	Met	Leu	Leu	Gly
			500				505						510		
Lys	Met	Cys	Ala	Pro	Arg	Tyr	Pro	Ser	Gly	Gly	Ala	Trp	Asp	Pro	His
		515					520					525			
Pro	Gly	Gln	Gly	Phe	Gly	Glu	Ser	Met	Tyr	Leu	Leu	Ser	Asp	Lys	Ala
	530					535				540					
Thr	Ser	Pro	Leu	Ser	Leu	Asp	Ala	Gly	Leu	Gly	Gln	Ala	Pro	Trp	Ser
545				550					555						560
Asp	Leu	Leu	Leu	Trp	Ala	Leu	Leu	Leu	Asn	Arg	Ala	Gln	Met	Ala	Met
			565					570						575	
Tyr	Phe	Trp	Glu	Met	Gly	Ser	Asn	Ala	Val	Ser	Ser	Ala	Leu	Gly	Ala
		580					585						590		
Cys	Leu	Leu	Leu	Arg	Val	Met	Ala	Arg	Leu	Glu	Pro	Asp	Ala	Glu	Glu
		595				600						605			
Ala	Ala	Arg	Arg	Lys	Asp	Leu	Ala	Phe	Lys	Phe	Glu	Gly	Met	Gly	Val
	610					615				620					
Asp	Leu	Phe	Gly	Glu	Cys	Tyr	Arg	Ser	Ser	Glu	Val	Arg	Ala	Ala	Arg
625				630						635					640
Leu	Leu	Leu	Arg	Arg	Cys	Pro	Leu	Trp	Gly	Asp	Ala	Thr	Cys	Leu	Gln
			645					650						655	
Leu	Ala	Met	Gln	Ala	Asp	Ala	Arg	Ala	Phe	Phe	Ala	Gln	Asp	Gly	Val
		660				665							670		
Gln	Ser	Leu	Leu	Thr	Gln	Lys	Trp	Trp	Gly	Asp	Met	Ala	Ser	Thr	Thr
		675				680						685			
Pro	Ile	Trp	Ala	Leu	Val	Leu	Ala	Phe	Phe	Cys	Pro	Pro	Leu	Ile	Tyr
	690					695					700				
Thr	Arg	Leu	Ile	Thr	Phe	Arg	Lys	Ser	Glu	Glu	Glu	Pro	Thr	Arg	Glu
705					710					715					720

-48-

Glu	Leu	Glu	Phe	Asp	Met	Asp	Ser	Val	Ile	Asn	Gly	Glu	Gly	Pro	Val	725	730	735
Gly	Thr	Ala	Asp	Pro	Ala	Glu	Lys	Thr	Pro	Leu	Gly	Val	Pro	Arg	Gln	740	745	750
Ser	Gly	Arg	Pro	Gly	Cys	Cys	Gly	Gly	Arg	Cys	Gly	Gly	Arg	Arg	Cys	755	760	765
Leu	Arg	Arg	Trp	Phe	His	Phe	Trp	Gly	Ala	Pro	Val	Thr	Ile	Phe	Met	770	775	780
Gly	Asn	Val	Val	Ser	Tyr	Leu	Leu	Phe	Leu	Leu	Leu	Phe	Ser	Arg	Val	785	790	795
Leu	Leu	Val	Asp	Phe	Gln	Pro	Ala	Pro	Pro	Gly	Ser	Leu	Glu	Leu	Leu	805	810	815
Leu	Tyr	Phe	Trp	Ala	Phe	Thr	Leu	Leu	Cys	Glu	Glu	Leu	Arg	Gln	Gly	820	825	830
Leu	Ser	Gly	Gly	Gly	Gly	Ser	Leu	Ala	Ser	Gly	Gly	Pro	Gly	Pro	Gly	835	840	845
His	Ala	Ser	Leu	Ser	Gln	Arg	Leu	Arg	Leu	Tyr	Leu	Ala	Asp	Ser	Trp	850	855	860
Asn	Gln	Cys	Asp	Leu	Val	Ala	Leu	Thr	Cys	Phe	Leu	Leu	Gly	Val	Gly	865	870	875
Cys	Arg	Leu	Thr	Pro	Gly	Leu	Tyr	His	Leu	Gly	Arg	Thr	Val	Leu	Cys	885	890	895
Ile	Asp	Phe	Met	Val	Phe	Thr	Val	Arg	Leu	Leu	His	Ile	Phe	Thr	Val	900	905	910
Asn	Lys	Gln	Leu	Gly	Pro	Lys	Ile	Val	Ile	Val	Ser	Lys	Met	Met	Lys	915	920	925
Asp	Val	Phe	Phe	Phe	Leu	Phe	Phe	Leu	Gly	Val	Trp	Leu	Val	Ala	Tyr	930	935	940
Gly	Val	Ala	Thr	Glu	Gly	Leu	Leu	Arg	Pro	Arg	Asp	Ser	Asp	Phe	Pro	945	950	955
Ser	Ile	Leu	Arg	Arg	Val	Phe	Tyr	Arg	Pro	Tyr	Leu	Gln	Ile	Phe	Gly	965	970	975
Gln	Ile	Pro	Gln	Glu	Asp	Met	Asp	Val	Ala	Leu	Met	Glu	His	Ser	Asn	980	985	990
Cys	Ser	Ser	Glu	Pro	Gly	Phe	Trp	Ala	His	Pro	Pro	Gly	Ala	Gln	Ala	995	1000	1005
Gly	Thr	Cys	Val	Ser	Gln	Tyr	Ala	Asn	Trp	Leu	Val	Val	Leu	Leu	Leu	1010	1015	1020
Val	Ile	Phe	Leu	Leu	Val	Ala	Asn	Ile	Leu	Leu	Val	Asn	Leu	Leu	Ile	1025	1030	1035
Ala	Met	Phe	Ser	Tyr	Thr	Phe	Gly	Lys	Val	Gln	Gly	Asn	Ser	Asp	Leu	1045	1050	1055
Tyr	Trp	Lys	Ala	Gln	Arg	Tyr	Arg	Leu	Ile	Arg	Glu	Phe	His	Ser	Arg	1060	1065	1070
Pro	Ala	Leu	Ala	Pro	Pro	Phe	Ile	Val	Ile	Ser	His	Leu	Arg	Leu	Leu	1075	1080	1085
Leu	Arg	Gln	Leu	Cys	Arg	Arg	Pro	Arg	Ser	Pro	Gln	Pro	Ser	Ser	Pro	1090	1095	1100
Ala	Leu	Glu	His	Phe	Arg	Val	Tyr	Leu	Ser	Lys	Glu	Ala	Glu	Arg	Lys	1105	1110	1115
Leu	Leu	Thr	Trp	Glu	Ser	Val	His	Lys	Glu	Asn	Phe	Leu	Leu	Ala	Arg	1125	1130	1135
Ala	Arg	Asp	Lys	Arg	Glu	Ser	Asp	Ser	Glu	Arg	Leu	Lys	Arg	Thr	Ser	1140	1145	1150
Gln	Lys	Val	Asp	Leu	Ala	Leu	Lys	Gln	Leu	Gly	His	Ile	Arg	Glu	Tyr	1155	1160	1165
Glu	Gln	Arg	Leu	Lys	Val	Leu	Glu	Arg	Glu	Val	Gln	Gln	Cys	Ser	Arg	1170	1175	1180
Val	Leu	Gly	Trp	Val	Ala	Glu	Ala	Leu	Ser	Arg	Ser	Ala	Leu	Leu	Pro	1185	1190	1195

Pro Gly Gly Pro Pro Pro Pro Asp Leu Pro Gly Ser Lys Asp
1205 1210

<210> 31
<211> 4646
<212> DNA
<213> Homo Sapiens

<400> 31
tcgaccacag cgtccgccca cgcgtccgcc cagcgcgtcg cccacgcgtc cgcccacgcg 60
tccgccacag cgtccggggg gaaagmramy cmygcktsms aaaaaccgtc acttaggaaa 120
agatgtcctt tcgggcagcc aggtcagca tgaggaaacag aaggaaatgac actctggaca 180
gcacccggac cctgtactcc agcgcgtctc ggagcacaga cttgtcttac agtgaaagcg 240
acttggtgaa ttttattcaa gcaaatttta agaaacgaga atgtgtcttc tttaccaaag 300
attccaaggc cagggagaat gtgtgcaagt gtggctatgc ccagagccag cacatggaag 360
gcaccagat caaccaaagt gagaaatgga actacaagaa acacaccaag gaatttccta 420
ccgacgcctt tggggatatt cagtttgaga cactggggaa gaaagggaag tatatacgtc 480
tgctctgcga cacggacgcg gaaatccttt acgagctgct gaccagcac tggcacctga 540
aaacacccaa cctgggtcatt tctgtgaccg ggggcgcaa gaacttcgcc ctgaagccgc 600
gcatgcgcaa gatcttcagc cggctcatct acatcgcgca gtccaaaggt gcttggttc 660
tcacgggagg caccattat ggctgatga agtacctcgg ggaggtggtg agagataaca 720
ccatcagcag gagttcagag gagaatattg tggccattgg catagcagct tggggcatgg 780
tctccaaccg ggacaccctc atcaggaatt gcgatgctga gggctatctt ttagcccaag 840
accttatgga tgacttcaca agagatccac tgtgtatcct ggacaacaac cacacacatt 900
tgctgctcgt ggacaatggc tgtcatggac atcccactgt cgaagcaaag ctccggaatc 960
agctagagaa gtatatctct gagcgacta ttcaagattc caactatggt ggcaagatcc 1020
ccattgtgtg ttttgcccaa ggaggtggaa aagagacttt gaaagccatc aatacctcca 1080
tcaaaaataa aattccttgt gtggtggtgg aaggctcggg ccagatcgct gatgtgatcg 1140
ctagcctggt ggaggtggag gatgccctga catcttctgc cgtcaaggag aagctggtgc 1200
gctttttacc ccgcacgggt tcccggctgc ctgaggagga gactgagagt tggatcaaatt 1260
ggctcaaaga aattctcgaa tgttctcacc tattaacagt tattaaaatg gaagaagctg 1320
gggatgaaat tgtgagcaat gccatctcct acgctctata caaagccttc agcaccagtg 1380
agcaagacaa ggataactgg aatgggcagc tgaagcttct gctggagtgg aaccagctgg 1440
acttagccaa tgatgagatt ttcaccaatg accgccgatg ggagtctgct gaccttcaag 1500
aagtcatggt tacggctctc ataaaggaca gacccaagtt tgtccgcctc tttctggaga 1560
atggcttgaa cctacggaag tttctcacc atgatgtcct cactgaactc tttccaacc 1620
acttcagcac gcttgtgtac cggaaatctgc agatcgccaa gaattcctat aatgatgccc 1680
tcctcacggt tgtctggaaa ctgggtgcga acttccgaag aggcttccgg aaggagaca 1740
gaaatggcgg ggacgagatg gacatagaac tccacagct gtctctatt actcggcacc 1800
ccttcaaagc tctcttcac tgggccattc ttcagaataa gaaggaaact tccaaagtca 1860
tttgggagca gaccaggggc tgcactctgg cagccctggg agccagcaag cttctgaaga 1920
ctctggccaa agtgaagaac gacatcaatg ctgctgggga gtccgaggag ctggctaattg 1980
agtacgagac ccgggctggt gagctgttca ctgagtggtt cagcagcgat gaagacttgg 2040
cagaacagct gctggtctat tcctgtgaag cttgggggtg aagcaactgt ctggagctgg 2100
cggtggaggg cacagaccag catttcacgc cccagcctgg ggtccagaat tttctttcta 2160
agcaatggta tggagagatt tcccagaca ccaagaactg gaagattatc ctgtgtctgt 2220
ttattatacc cttggtgggc tatgtggcgt tatcathtag gaagaaacct gtcgacaagc 2280
acaagaagct gctttggtac tatgtggcgt tcttcacctc tgcgtgttgc ctactgtctg ctcatggatt 2340
ggaatgtggt cttctacatc gccttcctcc tgcgtgttgc ctactgtctg ctcatggatt 2400
tccattcggg gccacacccc cccgagctgg tcctgtactc gctggtcttt gtcctcttct 2460
gtgatgaagt gagacagtgg tacgtaaag ggggtgaatta ttttactgac ctgtggaatg 2520
tgatggacac gctggggctt ttttacttca tagcaggaat tgtatttcgg ctccactctt 2580
ctaataaaaag ctctttgtat tctggacgag tcattttctg tctggactac attattttca 2640
ctctaagatt gatccacatt tttactgtaa gcagaaactt aggacccaag attataatgc 2700
tgcagaggat gctgatcgat gtgttcttct tcctgttctc ctttgcggtg tggatggtgg 2760
cctttggcgt ggccaggcaa gggatcctta ggcagaatga gcagcgctgg aggtggatat 2820
tccggtcggg catctacgag ccctacctgg ccagtgtcgg ccagggtccc agtgacgtgg 2880
atggtaccac gtatgacttt gccactgca ccttcaactg gaatgagtcc aagccactgt 2940
gtgtggagct ggatgagcac aacctgcccc ggttccccga gtggatcacc atccccctgg 3000
tgtgcatcta catgttatcc accaaccatc tgcgtgtcaa cctgctgggc gccatgtttg 3060

-50-

```

gctacacggt gggcaccgtc caggagaaca atgaccaggt ctggaagttc cagagggtact 3120
tcctgggtgca ggagtactgc agccgcctca atatcccctt ccccttcacg gtcttcgctt 3180
acttctacat ggtggggaag aagtgcctca agtggtgctg caaggagaaa aacatggagt 3240
cttctgtctg ctgtttcaaa aatgaagaca atgagactct ggcattggag ggtgtcatga 3300
aggaaaaacta ccttgtcaag atcaacacaa aagccaacga cacctcagag gaaatgaggc 3360
atcgatttag acaactggat acaaagctta atgatctcaa gggctctctg aaagagattg 3420
ctaataaaat caaataaaaac tgtatgaact ctaatggaga aaaatctaata tatagcaaga 3480
tcatattaag gaatgctgat gaacaatttt gctatcgact actaaatgag agattttcag 3540
accctgggt acatggtgga tgattttaaa tcaccctagt gtgctgagac cttgagaata 3600
aagtgtgtga ttggtttcat acttgaagac ggatataaag gaagaatatt tcctttatgt 3660
gtttctccag aatggtgcct gtttctctct gtgtctcaat gcctgggact ggagggtgat 3720
agttaaagtg tgttcttacc gcctcctttt tcctttaatc ttatttttga tgaacacata 3780
tataggagaa catctatcct atgaataaga acctggtcat gctttactcc tgtattgtta 3840
ttttgttcat ttccaattga ttctctactt ttcccttttt tgtattatgt gactaattag 3900
ttggcatatt gtwaaggatc tctcaaatga ggccagattc taaaacatgc tgcagcaaga 3960
ggaccccgct ctcttcagga aaagtgtttt catttctcag gatgcttctt acctgtcaga 4020
ggagggtgaca aggcagtctc ttgctctctt ggactcacca ggctcctatt gaaggacca 4080
ccccattcc taaatatgtg aaaagtcgcc caaaatgcaa ccttgaaagg cactactgac 4140
tttgtctta ttggatactc ctcttattta ttatttttcc attaaaaata atagctggct 4200
attatagaaa atttagacca tacagagatg tagaaagaac ataaattgtc cccattacct 4260
taaggtaatc actgctaaca atttctggat ggtttttcaa gtctattttt tttctatgta 4320
tgtctcaatt ctctttcaaa attttacaga atgttatcat actacatata tactttttat 4380
gtaagctttt tcacttagta ttttatcaaa tatgttttta ttatattcat agccttctta 4440
aacattatat caataattgc ataattggca acctctagcg attaccataa ttttgctcat 4500
tgaaggctat ctccagttga tcattgggat gagcatcitt gtgcatgaat cctattgctg 4560
tatttgggaa aattttccaa ggttagattc caataaatat ctatttatta ttaaaaaaaa 4620
aaaaaaaaag gcggccgctc tagagt 4646

```

<210> 32
 <211> 1104
 <212> PRT
 <213> Homo Sapiens

<400> 32

Met	Ser	Phe	Arg	Ala	Ala	Arg	Leu	Ser	Met	Arg	Asn	Arg	Arg	Asn	Asp
1				5					10					15	
Thr	Leu	Asp	Ser	Thr	Arg	Thr	Leu	Tyr	Ser	Ser	Ala	Ser	Arg	Ser	Thr
			20					25					30		
Asp	Leu	Ser	Tyr	Ser	Glu	Ser	Asp	Leu	Val	Asn	Phe	Ile	Gln	Ala	Asn
		35					40					45			
Phe	Lys	Lys	Arg	Glu	Cys	Val	Phe	Phe	Thr	Lys	Asp	Ser	Lys	Ala	Thr
	50					55					60				
Glu	Asn	Val	Cys	Lys	Cys	Gly	Tyr	Ala	Gln	Ser	Gln	His	Met	Glu	Gly
	65				70					75				80	
Thr	Gln	Ile	Asn	Gln	Ser	Glu	Lys	Trp	Asn	Tyr	Lys	Lys	His	Thr	Lys
			85						90					95	
Glu	Phe	Pro	Thr	Asp	Ala	Phe	Gly	Asp	Ile	Gln	Phe	Glu	Thr	Leu	Gly
			100					105					110		
Lys	Lys	Gly	Lys	Tyr	Ile	Arg	Leu	Ser	Cys	Asp	Thr	Asp	Ala	Glu	Ile
		115					120					125			
Leu	Tyr	Glu	Leu	Leu	Thr	Gln	His	Trp	His	Leu	Lys	Thr	Pro	Asn	Leu
	130					135					140				
Val	Ile	Ser	Val	Thr	Gly	Gly	Ala	Lys	Asn	Phe	Ala	Leu	Lys	Pro	Arg
	145				150					155				160	
Met	Arg	Lys	Ile	Phe	Ser	Arg	Leu	Ile	Tyr	Ile	Ala	Gln	Ser	Lys	Gly
			165						170					175	
Ala	Trp	Ile	Leu	Thr	Gly	Gly	Thr	His	Tyr	Gly	Leu	Met	Lys	Tyr	Ile
			180					185					190		
Gly	Glu	Val	Val	Arg	Asp	Asn	Thr	Ile	Ser	Arg	Ser	Ser	Glu	Glu	Asn
		195					200					205			

-51-

Ile	Val	Ala	Ile	Gly	Ile	Ala	Ala	Trp	Gly	Met	Val	Ser	Asn	Arg	Asp
	210					215					220				
Thr	Leu	Ile	Arg	Asn	Cys	Asp	Ala	Glu	Gly	Tyr	Phe	Leu	Ala	Gln	Tyr
225					230					235					240
Leu	Met	Asp	Asp	Phe	Thr	Arg	Asp	Pro	Leu	Cys	Ile	Leu	Asp	Asn	Asn
				245					250					255	
His	Thr	His	Leu	Leu	Val	Asp	Asn	Gly	Cys	His	Gly	His	Pro	Thr	
			260				265					270			
Val	Glu	Ala	Lys	Leu	Arg	Asn	Gln	Leu	Glu	Lys	Tyr	Ile	Ser	Glu	Arg
		275					280					285			
Thr	Ile	Gln	Asp	Ser	Asn	Tyr	Gly	Gly	Lys	Ile	Pro	Ile	Val	Cys	Phe
290					295					300					
Ala	Gln	Gly	Gly	Gly	Lys	Glu	Thr	Leu	Lys	Ala	Ile	Asn	Thr	Ser	Ile
305					310					315					320
Lys	Asn	Lys	Ile	Pro	Cys	Val	Val	Val	Glu	Gly	Ser	Gly	Gln	Ile	Ala
				325					330					335	
Asp	Val	Ile	Ala	Ser	Leu	Val	Glu	Val	Glu	Asp	Ala	Leu	Thr	Ser	Ser
			340					345					350		
Ala	Val	Lys	Glu	Lys	Leu	Val	Arg	Phe	Leu	Pro	Arg	Thr	Val	Ser	Arg
		355					360					365			
Leu	Pro	Glu	Glu	Glu	Thr	Glu	Ser	Trp	Ile	Lys	Trp	Leu	Lys	Glu	Ile
	370					375					380				
Leu	Glu	Cys	Ser	His	Leu	Leu	Thr	Val	Ile	Lys	Met	Glu	Glu	Ala	Gly
385					390					395					400
Asp	Glu	Ile	Val	Ser	Asn	Ala	Ile	Ser	Tyr	Ala	Leu	Tyr	Lys	Ala	Phe
			405						410					415	
Ser	Thr	Ser	Glu	Gln	Asp	Lys	Asp	Asn	Trp	Asn	Gly	Gln	Leu	Lys	Leu
			420					425					430		
Leu	Leu	Glu	Trp	Asn	Gln	Leu	Asp	Leu	Ala	Asn	Asp	Glu	Ile	Phe	Thr
		435					440					445			
Asn	Asp	Arg	Arg	Trp	Glu	Ser	Ala	Asp	Leu	Gln	Glu	Val	Met	Phe	Thr
	450					455					460				
Ala	Leu	Ile	Lys	Asp	Arg	Pro	Lys	Phe	Val	Arg	Leu	Phe	Leu	Glu	Asn
465					470					475					480
Gly	Leu	Asn	Leu	Arg	Lys	Phe	Leu	Thr	His	Asp	Val	Leu	Thr	Glu	Leu
			485						490					495	
Phe	Ser	Asn	His	Phe	Ser	Thr	Leu	Val	Tyr	Arg	Asn	Leu	Gln	Ile	Ala
			500					505					510		
Lys	Asn	Ser	Tyr	Asn	Asp	Ala	Leu	Leu	Thr	Phe	Val	Trp	Lys	Leu	Val
		515					520					525			
Ala	Asn	Phe	Arg	Arg	Gly	Phe	Arg	Lys	Glu	Asp	Arg	Asn	Gly	Arg	Asp
	530					535					540				
Glu	Met	Asp	Ile	Glu	Leu	His	Asp	Val	Ser	Pro	Ile	Thr	Arg	His	Pro
545					550					555					560
Leu	Gln	Ala	Leu	Phe	Ile	Trp	Ala	Ile	Leu	Gln	Asn	Lys	Lys	Glu	Leu
			565						570					575	
Ser	Lys	Val	Ile	Trp	Glu	Gln	Thr	Arg	Gly	Cys	Thr	Leu	Ala	Ala	Leu
			580					585					590		
Gly	Ala	Ser	Lys	Leu	Leu	Lys	Thr	Leu	Ala	Lys	Val	Lys	Asn	Asp	Ile
		595					600					605			
Asn	Ala	Ala	Gly	Glu	Ser	Glu	Glu	Leu	Ala	Asn	Glu	Tyr	Glu	Thr	Arg
	610					615					620				
Ala	Val	Glu	Leu	Phe	Thr	Glu	Cys	Tyr	Ser	Ser	Asp	Glu	Asp	Leu	Ala
625					630					635					640
Glu	Gln	Leu	Leu	Val	Tyr	Ser	Cys	Glu	Ala	Trp	Gly	Gly	Ser	Asn	Cys
			645						650					655	
Leu	Glu	Leu	Ala	Val	Glu	Ala	Thr	Asp	Gln	His	Phe	Ile	Ala	Gln	Pro
			660					665					670		
Gly	Val	Gln	Asn	Phe	Leu	Ser	Lys	Gln	Trp	Tyr	Gly	Glu	Ile	Ser	Arg
		675					680					685			

-52-

Asp	Thr	Lys	Asn	Trp	Lys	Ile	Ile	Leu	Cys	Leu	Phe	Ile	Ile	Pro	Leu
690						695					700				
Val	Gly	Cys	Gly	Phe	Val	Ser	Phe	Arg	Lys	Lys	Pro	Val	Asp	Lys	His
705					710					715					720
Lys	Lys	Leu	Leu	Trp	Tyr	Tyr	Val	Ala	Phe	Phe	Thr	Ser	Pro	Phe	Val
				725					730						735
Val	Phe	Ser	Trp	Asn	Val	Val	Phe	Tyr	Ile	Ala	Phe	Leu	Leu	Leu	Phe
			740					745					750		
Ala	Tyr	Val	Leu	Leu	Met	Asp	Phe	His	Ser	Val	Pro	His	Pro	Pro	Glu
		755					760					765			
Leu	Val	Leu	Tyr	Ser	Leu	Val	Phe	Val	Leu	Phe	Cys	Asp	Glu	Val	Arg
770						775					780				
Gln	Trp	Tyr	Val	Asn	Gly	Val	Asn	Tyr	Phe	Thr	Asp	Leu	Trp	Asn	Val
785					790					795					800
Met	Asp	Thr	Leu	Gly	Leu	Phe	Tyr	Phe	Ile	Ala	Gly	Ile	Val	Phe	Arg
				805					810						815
Leu	His	Ser	Ser	Asn	Lys	Ser	Ser	Leu	Tyr	Ser	Gly	Arg	Val	Ile	Phe
				820					825				830		
Cys	Leu	Asp	Tyr	Ile	Ile	Phe	Thr	Leu	Arg	Leu	Ile	His	Ile	Phe	Thr
		835					840					845			
Val	Ser	Arg	Asn	Leu	Gly	Pro	Lys	Ile	Ile	Met	Leu	Gln	Arg	Met	Leu
		850				855					860				
Ile	Asp	Val	Phe	Phe	Phe	Leu	Phe	Leu	Phe	Ala	Val	Trp	Met	Val	Ala
865					870					875					880
Phe	Gly	Val	Ala	Arg	Gln	Gly	Ile	Leu	Arg	Gln	Asn	Glu	Gln	Arg	Trp
				885					890						895
Arg	Trp	Ile	Phe	Arg	Ser	Val	Ile	Tyr	Glu	Pro	Tyr	Leu	Ala	Met	Phe
			900					905					910		
Gly	Gln	Val	Pro	Ser	Asp	Val	Asp	Gly	Thr	Thr	Tyr	Asp	Phe	Ala	His
		915					920					925			
Cys	Thr	Phe	Thr	Gly	Asn	Glu	Ser	Lys	Pro	Leu	Cys	Val	Glu	Leu	Asp
930					935						940				
Glu	His	Asn	Leu	Pro	Arg	Phe	Pro	Glu	Trp	Ile	Thr	Ile	Pro	Leu	Val
945					950					955					960
Cys	Ile	Tyr	Met	Leu	Ser	Thr	Asn	Ile	Leu	Leu	Val	Asn	Leu	Leu	Val
				965					970						975
Ala	Met	Phe	Gly	Tyr	Thr	Val	Gly	Thr	Val	Gln	Glu	Asn	Asn	Asp	Gln
			980					985					990		
Val	Trp	Lys	Phe	Gln	Arg	Tyr	Phe	Leu	Val	Gln	Glu	Tyr	Cys	Ser	Arg
		995					1000					1005			
Leu	Asn	Ile	Pro	Phe	Pro	Phe	Ile	Val	Phe	Ala	Tyr	Phe	Tyr	Met	Val
						1015					1020				
Val	Lys	Lys	Cys	Phe	Lys	Cys	Cys	Cys	Lys	Glu	Lys	Asn	Met	Glu	Ser
1025					1030					1035					104
Ser	Val	Cys	Cys	Phe	Lys	Asn	Glu	Asp	Asn	Glu	Thr	Leu	Ala	Trp	Glu
				1045					1050						1055
Gly	Val	Met	Lys	Glu	Asn	Tyr	Leu	Val	Lys	Ile	Asn	Thr	Lys	Ala	Asn
			1060					1065					1070		
Asp	Thr	Ser	Glu	Glu	Met	Arg	His	Arg	Phe	Arg	Gln	Leu	Asp	Thr	Lys
		1075					1080					1085			
Leu	Asn	Asp	Leu	Lys	Gly	Leu	Leu	Lys	Glu	Ile	Ala	Asn	Lys	Ile	Lys
1090							1095					1100			

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/29996

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C07K14/705 C12N15/12 C12Q1/68 C12N5/10 C07K16/28 G01N33/53 A61K38/17		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K C12Q A61K G01N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) BIOSIS, EPO-Internal, WPI Data, PAJ, MEDLINE, SCISEARCH, EMBASE, BIOTECHNOLOGY ABS, CHEM ABS Data, STRAND, GENSEQ, EMBL		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE GENEMBL 'Online! 16 February 1998 (1998-02-16) STRAUSBERG, R.: "ob70f05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1336737 3', mRNA sequence" XP002138823 Accession AA809355	1,2, 6-19, 25-35
X	DATABASE GENEMBL 'Online! 10 July 1998 (1998-07-10) MARRA ET AL.: "ub28d10.r1 Soares 2NbMT Mus musculus cDNA clone IMAGE:1379059 5' mRNA sequence" XP002149803 Accession AI050262	1,6-19, 25-35
-/--		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "&" document member of the same patent family		
Date of the actual completion of the international search 16 October 2000		Date of mailing of the international search report 30. 10. 00
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer ALCONADA RODRIG., A

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/29996

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE GENEMBL 'Online! 19 July 1997 (1997-07-19) STRAUSBERG, R.: "ni64e11.s1 NCI_CGAP_Pr12 Homo sapiens cDNA clone IMAGE:981644 mRNA sequence" XP002148641 Accession AA523749	1,3, 10-19, 25-35
X	WO 98 15657 A (ABBOTT LAB) 16 April 1998 (1998-04-16) page 4, line 7 -page 5, line 13 page 5, line 24 -page 7, line 28	1,4, 6-19, 25-35
Y	SEQ ID NOs. 9 and 25	20-24
X	WO 98 37093 A (CORIXA CORP) 27 August 1998 (1998-08-27) page 7, paragraph 2 page 9, paragraphs 2,3 page 13 -page 17 page 21, paragraph 3	1,4, 6-19, 25-35
Y	SEQ ID NOs: 109 and 112	20-24
X	DATABASE GENEMBL 'Online! 18 November 1997 (1997-11-18) STRAUSBERG, R.: "nt76b07.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1204405, mRNA" XP002148642 Accession AA654650	1,5-19, 25-35
Y		20-24
Y	DATABASE GENEMBL 'Online! 30 November 1998 (1998-11-30) SHIMIZU, N.: "Homo sapiens mRNA complete cds." XP002148643 Accession number AB001535 -& NAGAMINE ET AL.: "Molecular cloning of a novel putative Ca2+ channel protein (TRPC7) highly expressed in brain" GENOMICS, vol. 54, 15 November 1998 (1998-11-15), pages 124-131, XP000938744 the whole document	20-24

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/29996

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ZHU, XI ET AL: "Molecular cloning of a widely expressed human homologue for the <i>Drosophila trp</i> gene." FEBS LETTERS, (1995) VOL. 373, NO. 3, PP. 193-198., XP000907241 page 194; figures 1,3 ---	20,21, 23,25, 26,28, 29,31
A	HUNTER JOHN J ET AL: "Chromosomal localization and genomic characterization of the mouse melastatin gene (<i>Mln1</i>)." GENOMICS NOV. 15, 1998, vol. 54, no. 1, 15 November 1998 (1998-11-15), pages 116-123, XP000910696 ISSN: 0888-7543 cited in the application page 119; figure 2 ---	20,21,23
A	WES PAUL D ET AL: "TRPC1, a human homolog of a <i>Drosophila</i> store-operated channel." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 1995, vol. 92, no. 21, 1995, pages 9652-9656, XP002138820 ISSN: 0027-8424 the whole document ---	20,21, 23,25, 26,28, 29,31
A	ZHU, XI ET AL: "Trp, A novel mammalian gene family essential for agonist-activated capacitative Ca-2+ entry." CELL, vol. 85, no. 5, 1996, pages 661-671, XP000907242 page 662 page 665 figures 1,5,6 ---	20,21, 25,26, 28,29,31
A	GARCIA REYNALDO L ET AL: "Differential expression of mammalian TRP homologues across tissues and cell lines." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS 1997, vol. 239, no. 1, 1997, pages 279-283, XP002138822 ISSN: 0006-291X See Materials and Methods figure 1 --- -/-	25,26, 28-30

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/29996

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SINKINS WILLIAM G ET AL: "Functional expression of TrpC1: A human homologue of the Drosophila Trp channel." BIOCHEMICAL JOURNAL APRIL, 1998, vol. 331, no. 1, April 1998 (1998-04), pages 331-339, XP000864583 ISSN: 0264-6021 page 333-335; figures 3-5	24
A	PREUSS KLAUS-DIETER ET AL: "Expression and characterization of a trp1 homolog from rat." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS NOV. 7, 1997, vol. 240, no. 1, 7 November 1997 (1997-11-07), pages 167-172, XP002138821 ISSN: 0006-291X figure 2	24
A	OBUKHOV, ALEXANDER G. ET AL: "Direct activation of trp1 cation channels by G-alpha-11 subunits." EMBO (EUROPEAN MOLECULAR BIOLOGY ORGANIZATION) JOURNAL, (1996) VOL. 15, NO. 21, PP. 5833-5838., XP000907243 figure 2	24
P,X	WO 99 09199 A (RYAZANOV ALEXEY G ;PAVUR KAREN S (US); HAIT WILLIAM N (US); UNIV M) 25 February 1999 (1999-02-25) see melanome kinase polynucleotide and polypeptide sequences on page 16-17	1,3, 10-19, 25-36
P,X	WO 99 09166 A (SHAPERO MICHAEL H ;DENDREON CORP (US); LAUS REINER (US); TSAVALER) 25 February 1999 (1999-02-25) page 17, line 24 -page 18, line 9 page 25, line 19-32 page 28, line 1-4 SEQ ID NOs: 27, 28 and 31.	1,5-19, 25-35
T	SCHARENBERG A M ET AL: "MLSN-1/SOC-1 defines a widely expressed Ca2+/cation channel family involved in Ca2+ homeostasis and store-operated Ca2+ signaling." FIFTY-THIRD ANNUAL MEETING OF THE SOCIETY OF GENERAL PHYSIOLOGISTS;WOODS HOLE, MASSACHUSETTS, USA; SEPTEMBER 9-11, 1999, vol. 114, no. 1, July 1999 (1999-07), page 14a XP000910708 Journal of General Physiology July, 1999 ISSN: 0022-1295	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/29996

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

As a result of the prior review under R. 40.2(e) PCT,
no additional fees are to be refunded.

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☒ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
1-36
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☒ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Present claims 1-5, 10-13, 16-19, 32-35 relate to an extremely large number of possible polynucleotides, polypeptides encoded by them, binding polypeptides, and kits and pharmaceutical compositions containing said polypeptides and polynucleotides. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the compounds claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to the polynucleotide of SEQ ID NOs: 1, 27, 29 and 31 and the corresponding polypeptide of SEQ ID NOs: 2, 28, 30 and 32.

Present claims 16 and 17 relate to an extremely large number of possible compounds, namely, a polypeptide that binds to the polypeptide of the invention. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the compounds claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to an antibody, antibody fragment, F(ab)₂ fragment or a fragment including a CDR3 region selective for the polypeptides of the invention.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1, 6-36 (partially) and 2 (complete)

An isolated nucleic acid molecule comprising a nucleic acid molecule that hybridizes to a nucleic acid molecule of SEQ ID NO:1 and which code for a SOC/CRAC polypeptide, nucleic acid molecules that differ in codon sequence due to degeneracy of the genetic code and complement thereof, polynucleotides which are not identical to the SEQ ID or sequences of GenBank accession number of Table 1; expression vector, host cells; polypeptide encoded thereof (SEQ ID NO:2); polypeptides binding to the polypeptide of SEQ ID NO:2, including antibodies; kits comprising agents that selectively bind to the polynucleotide (SEQ ID NO:1) or polypeptide (SEQ ID NO:2) of the invention; pharmaceutical compositions containing the polynucleotide or polypeptides of the invention; a method for isolating the SOC/CRAC molecule having SOC/CRAC calcium channel activity comprising contacting a binding molecule that is SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing SOC/CRAC molecules allowing the formation of the complex, detecting the formation of the complex, isolating the SOC/CRAC molecule and determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity; a method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity; a method to determine the level of SOC/CRAC expression in a subject, including expression of SOC/CRAC polypeptide or mRNA in a tissue or biological fluid sample using PCR, Northern blotting, and mono- and polyclonal antisera and a method for identifying agents useful in the modulation of the SOC/CRAC polypeptide kinase activity, comprising the use of aminoacids 999-1180 from SEQ ID NO:4 as a candidate kinase.

2. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:3 and to the encoded polypeptide of SEQ ID NO:4

3. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:5 and to the encoded polypeptide of SEQ ID NO:6

4. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:7 and to the encoded polypeptide of SEQ ID NO:8

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

5. Claims: 1,6-36 (partially) and 37 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:23 and to the encoded polypeptide of SEQ ID NO:24

6. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:25 and to the encoded polypeptide of SEQ ID NO:26

7. Claims: 1,10-36 (partially) and 3 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:27 and to the encoded polypeptide of SEQ ID NO:28

8. Claims: 1,6-36 (partially) and 4 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:29 and to the encoded polypeptide of SEQ ID NO:30

9. Claims: 1,6-36 (partially) and 5 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:31 and to the encoded polypeptide of SEQ ID NO:32.

INTERNATIONAL SEARCH REPORT

Information on patent family members

In ternational Application No

PCT/US 99/29996

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9815657 A	16-04-1998	US 5919638 A EP 0954599 A US 6110675 A	06-07-1999 10-11-1999 29-08-2000
WO 9837093 A	27-08-1998	AU 6181898 A CN 1252837 T EP 1005546 A NO 994069 A PL 335348 A ZA 9801585 A	09-09-1998 10-05-2000 07-06-2000 22-10-1999 25-04-2000 04-09-1998
WO 9909199 A	25-02-1999	AU 9110098 A	08-03-1999
WO 9909166 A	25-02-1999	AU 9021898 A EP 1005549 A	08-03-1999 07-06-2000

CORRECTED VERSION

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
13 July 2000 (13.07.2000)

PCT

(10) International Publication Number
WO 00/40614 A3

(51) International Patent Classification⁷: C07K 14/705,
C12N 15/12, C12Q 1/68, C12N 5/10, C07K 16/28, G01N
33/53, A61K 38/17

(74) Agent: PLUMER, Elizabeth, R.; Wolf, Greenfield &
Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).

(21) International Application Number: PCT/US99/29996

(81) Designated States (*national*): AU, CA, JP, US.

(22) International Filing Date:
20 December 1999 (20.12.1999)

(84) Designated States (*regional*): European patent (AT, BE,
CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC,
NL, PT, SE).

(25) Filing Language: English

Published:
— with international search report

(26) Publication Language: English

(30) Priority Data:
60/114,220 30 December 1998 (30.12.1998) US
60/120,018 29 January 1999 (29.01.1999) US
60/140,415 22 June 1999 (22.06.1999) US

(88) Date of publication of the international search report:
22 February 2001

(48) Date of publication of this corrected version:
30 August 2001

(71) Applicant (*for all designated States except US*): BETH
ISRAEL DEACONESS MEDICAL CENTER, INC.
[US/US]; 1 Deaconess Road, Boston, MA 02215 (US).

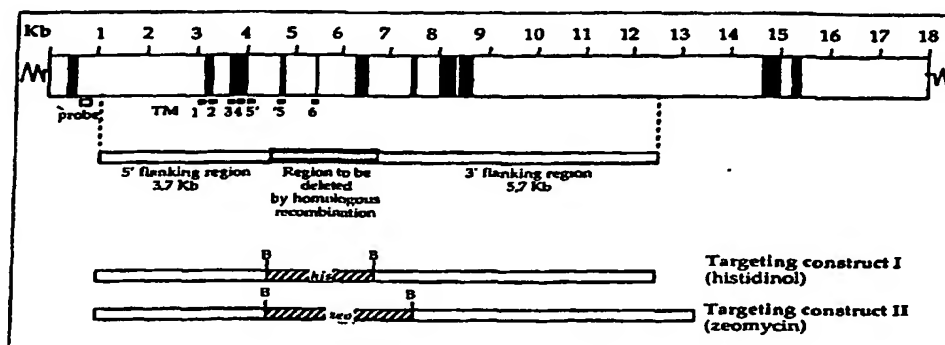
(15) Information about Correction:
see PCT Gazette No. 35/2001 of 30 August 2001, Section
II

(72) Inventor; and

(75) Inventor/Applicant (*for US only*): SCHARENBERG,
Andrew, M. [US/US]; 12 Skyview Road, Lexington, MA
02420 (US).

*For two-letter codes and other abbreviations, refer to the "Guid-
ance Notes on Codes and Abbreviations" appearing at the begin-
ning of each regular issue of the PCT Gazette.*

(54) Title: CHARACTERIZATION OF THE SOC/CRAC CALCIUM CHANNEL PROTEIN FAMILY



(57) Abstract: Nucleic acids encoding SOC/CRAC calcium channel polypeptides, including fragments and biologically functional variants thereof and encoded polypeptides are provided. The nucleic acids and polypeptides disclosed herein are useful as therapeutic and diagnostic agents. Agents that selectively bind to the foregoing polypeptides and genes also are provided.

CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY**Field of the Invention**

This invention relates to nucleic acids coding for a novel family of calcium channel polypeptides, the encoded polypeptides, unique fragments of the foregoing, and methods of making and using same.

Background of the Invention

Calcium channels are membrane-spanning, multi-subunit proteins that facilitate the controlled transport ("flux") of Ca^{2+} ions into and out of cells. Cells throughout the animal kingdom, and at least some bacterial, fungal and plant cells, possess one or more types of calcium channels. In general, "excitable" cells, such as neurons of the central nervous system, peripheral nerve cells, and muscle cells, including those of skeletal muscles, cardiac muscles, and venous and arterial smooth muscles, possess voltage-dependent calcium channels. In a voltage-dependent calcium channel, the transport of Ca^{2+} ions into and out of the cells requires a certain minimal level of depolarization (the difference in potential between the inside of the cell bearing the channel and the extracellular environment) with the rate of Ca^{2+} cell flux dependent on the difference in potential. In "non-excitable" cells, calcium influx is thought to occur predominantly in response to stimuli which cause the release of calcium from intracellular stores. This process, termed *store operated calcium influx*, is not well understood.

Characterization of a particular type of calcium channel by analysis of whole cells is complicated by the presence of mixed populations of different types of calcium channels in the majority of cells. Although single-channel recording methods can be used to examine individual calcium channels, such analysis does not reveal information related to the molecular structure or biochemical composition of the channel. Furthermore, in this type of analysis, the channel is isolated from other cellular constituents that might be important for the channel's natural functions and pharmacological interactions. To study the calcium channel structure-function relationship, large amounts of pure channel protein are needed. However, acquiring large amounts of pure protein is difficult in view of the complex nature of these multisubunit proteins, the varying concentrations of calcium channel proteins in tissue sources, the presence of mixed populations of calcium channel proteins in tissues, and the modifications of the native protein that can occur during the isolation procedure.

Summary of the Invention

The invention is based on the identification of a novel family of calcium channel polypeptides and the molecular cloning and partial characterization of a novel member of this family that is expressed predominantly in human hematopoietic cells, liver, and kidney. This newly identified family of calcium channel polypeptides is designated, "SOC" or "CRAC" or "ICRAC", for Store Operated Channels or Calcium Release Activated Channels. Although not wishing to be bound to any particular theory or mechanism, it is believed that the SOC/CRAC calcium channel polypeptides are transmembrane polypeptides that modulate Ca^{2+} flux "into" and "out of" a cell, for example, in certain instances they may be activated upon depletion of Ca^{2+} from intracellular calcium stores, allowing Ca^{2+} influx into the cell. Accordingly, the compositions disclosed herein are believed to be useful for modulating calcium transport into and out of such intracellular stores and for the treatment of disorders that are characterized by aberrant calcium transport into and out of such intracellular stores. In particular, we believe that the SOC/CRAC calcium channel polypeptides disclosed herein play an important role in the influx of extracellular calcium by mediating the refilling of intracellular calcium stores following their depletion. Accordingly, we believe that the compositions for expressing functional SOC/CRAC calcium channel polypeptides in cells, as disclosed herein, are useful for treating patients having conditions that are characterized by reduced extracellular calcium influx into their SOC/CRAC-expressing cells. Additionally, the compositions of the invention are useful for delivering therapeutic and/or imaging agents to cells which preferentially express SOC/CRAC calcium channel polypeptides and, in particular, for delivering such agents to hematopoietic cells, liver, heart, spleen, and kidney to modulate proliferation and growth of these cells. Moreover, in view of the importance of cellular calcium levels to cell viability, we believe that SOC-2/CRAC-1, SOC-3/CRAC-2, and SOC-4/CRAC-3 as disclosed herein, and/or other members of the SOC/CRAC family of calcium channel polypeptides, represent an ideal target for designing and/or identifying (e.g., from molecular libraries) small molecule inhibitors that block lymphocyte proliferation, as well as other binding agents that selectively bind to SOC/CRAC polypeptides to which drugs or toxins can be conjugated for delivery to SOC/CRAC polypeptide expressing cells.

The invention is based, in part, on the molecular cloning and sequence analysis of the novel SOC/CRAC calcium channel molecules disclosed herein (also referred to as a "SOC-2/CRAC-1 molecule," a "SOC-3/CRAC-2 molecule," and/or "SOC-4/CRAC-3 molecule") that are predominantly expressed in human hematopoietic cells, liver, spleen, heart, and

kidney (SOC-2/CRAC-1), kidney and colon (SOC-3/CRAC-2), and prostate (SOC-4/CRAC-3 molecule). As used herein, a "SOC/CRAC molecule" embraces a "SOC/CRAC calcium channel nucleic acid" (or "SOC/CRAC nucleic acid") and a "SOC/CRAC calcium channel polypeptide" (or "SOC/CRAC polypeptide"). Homologs and alleles also are embraced within the meaning of a SOC/CRAC calcium channel molecule.

According to one aspect of the invention, isolated SOC/CRAC nucleic acids which code for one or more member(s) of the SOC/CRAC family of calcium channel polypeptides or unique fragments thereof are provided. The isolated nucleic acids refer to one or more of the following:

(a) nucleic acid molecules which hybridize under stringent conditions to a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;

(b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;

(c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and

(d) complements of (a), (b) or (c).

The invention in another aspect provides an isolated nucleic acid molecule selected from the group consisting of (a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:29, and SEQ ID NO:31, (b) complements of (a), provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to any sequence selected from a sequence group consisting of (1) sequences having the SEQ. ID NOS. or GenBank accession numbers of Table I, (2) complements of (1), and (3) fragments of (1) and (2).

According to yet another aspect of the invention, isolated SOC/CRAC polypeptides are provided. The isolated SOC/CRAC polypeptide molecules are encoded by one or more SOC/CRAC nucleic acid molecules of the invention. Preferably, the SOC/CRAC polypeptide contains one or more polypeptides selected from the group consisting of the polypeptides having SEQ. ID Nos. 2, 4, 6, 8, 24, 26, 28, 30, and 32. In other embodiments, the isolated polypeptide may be a fragment or variant of the foregoing SOC/CRAC polypeptide molecules of sufficient length to represent a sequence unique within the human genome, and identifying

with a polypeptide that functions as a calcium channel, provided that the fragment excludes a sequence of contiguous amino acids identified in Table II, and/or excludes a sequence of contiguous amino acids encoded for by a nucleic acid sequence identified in Table I. In another embodiment, immunogenic fragments of the polypeptide molecules described above are provided.

According to another aspect of the invention, isolated SOC/CRAC binding agents (e.g., polypeptides) are provided which selectively bind to a SOC/CRAC molecule (e.g., a SOC/CRAC polypeptide encoded by the isolated nucleic acid molecules of the invention). Preferably, the isolated binding agents selectively bind to a polypeptide which comprises the sequence selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32, or unique fragments thereof. In the preferred embodiments, the isolated binding polypeptides include antibodies and fragments of antibodies (e.g., Fab, F(ab)₂, Fd and antibody fragments which include a CDR3 region which binds selectively to a SOC/CRAC polypeptide). Preferably, the antibodies for human therapeutic applications are human antibodies.

According to another aspect of the invention, a pharmaceutical composition containing a pharmaceutically effective amount of an isolated SOC/CRAC nucleic acid, an isolated SOC/CRAC polypeptide, or an isolated SOC/CRAC binding polypeptide in a pharmaceutically acceptable carrier also is provided. The pharmaceutical compositions are useful in accordance with therapeutic methods disclosed herein.

According to yet another aspect of the invention, a method for isolating a SOC/CRAC molecule is provided. The method involves:

a) contacting a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample that is believed to contain one or more SOC/CRAC molecules, under conditions to form a complex of the SOC/CRAC nucleic acid or the SOC/CRAC binding polypeptide and the SOC/CRAC molecule;

b) detecting the presence of the complex;

c) isolating the SOC/CRAC molecule from the complex; and

d) determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity. As used herein "SOC/CRAC calcium channel activity" refers to the transport of Ca²⁺ into and out of intracellular stores that is mediated by a SOC/CRAC

polypeptide. In general, the SOC/CRAC calcium channel activity is initiated by a reduction or depletion of intracellular calcium stores.

In certain embodiments, the SOC/CRAC nucleic acid is a SOC-2/CRAC-1 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 27, or complements thereof); in certain other
5 embodiments, the SOC/CRAC nucleic acid is a SOC-3/CRAC-2 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 29, or complements thereof); in further embodiments, the SOC/CRAC nucleic acid is a SOC-4/CRAC-3 nucleic acid (e.g., a nucleic acid having SEQ. ID NO. 31, or complements thereof). In yet other embodiments, the SOC/CRAC polypeptide is a SOC-2/CRAC-1 binding polypeptide (e.g., an antibody that selectively binds to a SOC-
10 2/CRAC-1 polypeptide). In yet further embodiments, the SOC/CRAC polypeptide is a SOC-3/CRAC-2 binding polypeptide (e.g., an antibody that selectively binds to a SOC-3/CRAC-2 polypeptide). In some embodiments, the SOC/CRAC polypeptide is a SOC-4/CRAC-3 binding polypeptide (e.g., an antibody that selectively binds to a SOC-4/CRAC-3 polypeptide). In the preferred embodiments, the isolated binding polypeptides include
15 antibodies and fragments of antibodies (e.g., Fab, F(ab)₂, Fd and antibody fragments which include a CDR3 region which binds selectively to a SOC-2/CRAC-1, to a SOC-3/CRAC-2, and/or to a SOC-4/CRAC-3 polypeptide). Preferably the isolated binding polypeptides or other binding agents selectively bind to a single SOC/CRAC molecule, i.e., are capable of distinguishing between different members of the SOC/CRAC family. Accordingly, one or
20 more SOC/CRAC binding agents can be contained in a single composition (e.g., a pharmaceutical composition) to identify multiple SOC/CRAC molecules *in vivo* or *in vitro*.

According to yet another aspect of the invention, a method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity is provided. The method involves:

- 25 a) contacting a SOC/CRAC polypeptide with a candidate agent suspected of modulating SOC/CRAC calcium channel activity, under conditions sufficient to allow the candidate agent to interact selectively with (e.g. bind to) the SOC/CRAC polypeptide;
- b) detecting a Ca²⁺ concentration of step (b) associated with the SOC/CRAC calcium channel activity of the SOC/CRAC polypeptide in the presence of the candidate agent; and
- 30 c) comparing the Ca²⁺ concentration of step (b) with a control Ca²⁺ concentration of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC calcium channel activity.

According to another aspect of the invention, a method for identifying agents useful in the modulation of a SOC/CRAC polypeptide kinase activity is provided. The method involves:

a) contacting a SOC/CRAC polypeptide with kinase activity with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity;

b) detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and

c) comparing the kinase activity of step (b) with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC kinase activity. In some embodiments the SOC/CRAC polypeptide comprises amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24), or a fragment thereof that retains the kinase activity.

According to yet another aspect of the invention, a method for determining the level of expression of a SOC/CRAC polypeptide in a subject is provided. The method involves:

a) measuring the expression of a SOC/CRAC polypeptide in a test sample, and

b) comparing the measured expression of the SOC/CRAC polypeptide in the test sample to the expression of a SOC/CRAC polypeptide in a control containing a known level of expression to determine the level of SOC/CRAC expression in the subject. Expression is defined as SOC/CRAC mRNA expression or SOC/CRAC polypeptide expression. Various methods can be used to measure expression. The preferred embodiments of the invention utilize PCR and Northern blotting for measuring mRNA expression, and monoclonal or polyclonal SOC/CRAC antisera as reagents for measuring SOC/CRAC polypeptide expression. In preferred embodiments, the SOC/CRAC molecule (nucleic acid and/or polypeptide) is SOC-2/CRAC-1. In other preferred embodiments, the SOC/CRAC molecule is SOC-3/CRAC-2. In yet further preferred embodiments, the SOC/CRAC molecule is SOC-4/CRAC-3. In certain embodiments, the test samples include biopsy samples and biological fluids such as blood. The method is useful, e.g., for assessing the presence or absence or stage of a proliferative disorder in a subject.

The invention also contemplates kits comprising a package including assays for SOC/CRAC epitopes, SOC/CRAC nucleic acids, and instructions, and optionally related materials such as controls, for example, a number, color chart, or an epitope of the expression product of the foregoing isolated nucleic acid molecules of the invention for comparing, for

example, the level of SOC/CRAC polypeptides or SOC/CRAC nucleic acid forms (wild-type or mutant) in a test sample to the level in a control sample having a known amount of a SOC/CRAC nucleic acid or SOC/CRAC polypeptide. This comparison can be used to assess in a subject a risk of developing a cancer or the progression of a cancer. The kits may also include assays for other known genes, and expression products thereof, associated with, for example, proliferative disorders (e.g., BRCA, p53, etc.). In a preferred embodiment, the kit comprises a package containing: (a) a binding agent that selectively binds to an isolated nucleic acid of the invention or an expression product thereof to obtain a measured test value, (b) a control containing a known amount of a SOC/CRAC nucleic acid or a SOC/CRAC polypeptide to obtain a measured control value, and (c) instructions for comparing the measured test value to the measured control value to determine the amount of SOC/CRAC nucleic acid or expression product thereof in a sample.

The invention provides isolated nucleic acid molecules, unique fragments thereof, expression vectors containing the foregoing, and host cells containing the foregoing. The invention also provides isolated binding polypeptides and binding agents which bind such polypeptides, including antibodies, and pharmaceutical compositions containing any of the compositions of the invention. The foregoing can be used, *inter alia*, in the diagnosis or treatment of conditions characterized by the aberrant expression levels and/or the presence of mutant forms of a SOC/CRAC nucleic acid or polypeptide. The invention also provides methods for identifying agents that alter the function of the SOC/CRAC polypeptide.

These and other aspects of the invention, as well as various advantages and utilities, will be more apparent with reference to the detailed description of the preferred embodiments.

Brief Description of the Sequences

SEQ ID NO:1 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:2 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:1).

SEQ ID NO:3 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:4 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:3).

SEQ ID NO:5 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:6 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:5).

SEQ ID NO:7 is a partial nucleotide sequence of the mouse homologue (mSOC-2/CRAC-1) of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:8 is the predicted amino acid sequence of the translation product of the mSOC-2/CRAC-1 cDNA (SEQ ID NO:7).

5 SEQ ID NO:9 is the nucleotide sequence of the mouse MLSN-1 (SOC-1) cDNA.

SEQ ID NO:10 is the predicted amino acid sequence of the translation product of the mouse MLSN-1 (SOC-1) cDNA (SEQ ID NO:9).

SEQ ID NO:11 is the nucleotide sequence of a human calcium channel cDNA with GenBank Acc. no.: AB001535.

10 SEQ ID NO:12 is the predicted amino acid sequence of the translation product of the human calcium channel cDNA with GenBank Acc. no.: AB001535 (SEQ ID NO:11).

SEQ ID NO:13 is the amino acid sequence of a *C. Elegans* polypeptide at the c05c12.3 locus.

15 SEQ ID NO:14 is the amino acid sequence of a *C. Elegans* polypeptide at the F54D1 locus.

SEQ ID NO:15 is the amino acid sequence of a *C. Elegans* polypeptide at the t01H8 locus.

SEQ ID NO:16 is the nucleotide sequence of a mouse kidney cDNA with GenBank Acc. no.: AI226731.

20 SEQ ID NO:17 is the predicted amino acid sequence of the translation product of the mouse kidney cDNA with GenBank Acc. no.: AI226731 (SEQ ID NO:16).

SEQ ID NO:18 is the nucleotide sequence of a human brain cDNA with GenBank Acc. no.: H18835.

25 SEQ ID NO:19 is the predicted amino acid sequence of the translation product of the human brain cDNA with GenBank Acc. no.: H18835 (SEQ ID NO:18).

SEQ ID NO:20 is the nucleotide sequence of the human EST with GenBank Acc. no.: AA419592.

SEQ ID NO:21 is the nucleotide sequence of the human EST with GenBank Acc. no.: AA419407.

30 SEQ ID NO:22 is the nucleotide sequence of the mouse EST with GenBank Acc. no.: AI098310.

SEQ ID NO:23 is a partial nucleotide sequence of the human SOC-2/CRAC-1 cDNA that contains the SOC-2/CRAC-1 sequences of SEQ ID NO:1, SEQ ID NO:3, and SEQ ID NO:5.

SEQ ID NO:24 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:23).

SEQ ID NO:25 is a partial nucleotide sequence of the human SOC-3/CRAC-2 cDNA.

SEQ ID NO:26 is the predicted amino acid sequence of the translation product of human SOC-3/CRAC-2 cDNA (SEQ ID NO:25).

SEQ ID NO:27 is the full nucleotide sequence of the human SOC-2/CRAC-1 cDNA.

SEQ ID NO:28 is the predicted amino acid sequence of the translation product of human SOC-2/CRAC-1 cDNA (SEQ ID NO:27).

SEQ ID NO:29 is the full nucleotide sequence of the human SOC-3/CRAC-2 cDNA.

SEQ ID NO:30 is the predicted amino acid sequence of the translation product of human SOC-3/CRAC-2 cDNA (SEQ ID NO:29).

SEQ ID NO:31 is the full nucleotide sequence of the human SOC-4/CRAC-3 cDNA.

SEQ ID NO:32 is the predicted amino acid sequence of the translation product of human SOC-4/CRAC-3 cDNA (SEQ ID NO:31).

Brief Description of the Drawings

Figure 1 is a schematic depicting the intron/exon organization of the chicken SOC-2/CRAC-1 genomic sequence, as well as the putative transmembrane (TM) domains, and the targeting constructs utilized in the knockout experiments.

Detailed Description of the Invention

One aspect of the invention involves the partial cloning of cDNAs encoding members of a novel family of calcium channel polypeptides, referred to herein as "SOC/CRAC" (designated "SOC" or "CRAC" or "ICRAC", for Sore Operated Channels or Calcium Release Activated Channels, or CECH). Although not intending to be bound to any particular mechanism or theory, we believe that a SOC/CRAC family member is a transmembrane calcium channel that modulates Ca^{2+} flux "into" and "out of" a cell; in certain instances it may be activated upon depletion of Ca^{2+} from intracellular calcium stores, allowing Ca^{2+} influx into the cell.

The first three isolated SOC/CRAC members disclosed herein, define a new family of calcium channels which is distinct from previously described calcium channels, such as voltage gated calcium channels, ryanodine receptor/inositol-1,4,5-triphosphate receptor

channels, and Transient Receptor Potential (TRP) channels. The SOC/CRAC family of calcium channels exhibits high selectivity (with a P_{Ca}/P_{Na} ratio near 1000), a unitary conductance below the detection level of the patch clamp method (the conductance estimated at approximately 0.2 picosiemens), and are subject to inhibition by high intracellular calcium levels. Although not intending to be bound to any particular mechanism or theory, we believe that SOC/CRAC calcium channels are responsible for the majority of, for example, calcium entry which occurs when intracellular calcium stores are depleted, and that SOC/CRAC currents are important for initiating various types of calcium-dependent processes. Thus, we believe that SOC/CRAC calcium channels play an important role in cellular calcium homeostasis by, e.g., modulating the supply of calcium to refill intracellular stores when depleted.

The isolated full-length sequence of a representative, first member of the SOC/CRAC family, human SOC/CRAC nucleic acid (cDNA), SOC-2/CRAC-1, is represented as the nucleic acid of SEQ ID NO:27. This nucleic acid sequence codes for the SOC-2/CRAC-1 polypeptide with the predicted amino acid sequence disclosed herein as SEQ ID NO:28. A homologous mouse cDNA sequence (>90% identity to the human at the nucleotide level) is represented as the nucleic acid of SEQ ID NO:7, and codes for a unique fragment of a mouse SOC-2/CRAC-1 polypeptide having the predicted, partial amino acid sequence represented as SEQ ID NO:8. Analysis of the SOC-2/CRAC-1 partial sequence by comparison to nucleic acid and protein databases show that SOC-2/CRAC-1 shares a limited homology to mouse MLSN-1 (SOC-1, SEQ ID NOs: 9 and 10). Limited homology is also shared between SOC-2/CRAC-1 and three *C. Elegans* polypeptides (SEQ ID NOs: 13, 14, and 15). We further believe that SOC-2/CRAC-1 plays a role in the regulation of cellular Ca^{2+} fluxing and, in particular, lymphocyte Ca^{2+} fluxing.

A second member of the human SOC/CRAC family of calcium channels, SOC-3/CRAC-2, is represented as the nucleic acid of SEQ ID NO:29, and codes for the human SOC-3/CRAC-2 polypeptide having the predicted amino acid sequence represented as SEQ ID NO:30 (this molecule may also be referred to as CECH2). SOC-3/CRAC-2 is predominantly expressed in human hematopoietic cells (including peripheral blood lymphocytes, liver, bone marrow, spleen, thymus, lymph nodes, heart, and kidney. Expression can also be detected (at lesser levels) in brain, skeletal muscle colon, small intestine, placenta, lung, and cells (cell lines) such as HL-60, HeLa, K562, MOLT-4, SW-480, A459, and G361.

A third member of the human SOC/CRAC family of calcium channels, SOC-4/CRAC-3, is represented as the nucleic acid of SEQ ID NO:31, and codes for the human SOC-4/CRAC-3 polypeptide having the predicted amino acid sequence represented as SEQ ID NO:32 (this molecule may also be referred to as CECH6). It specifically expressed in the prostate gland/cells.

As used herein, a SOC/CRAC calcium channel nucleic acid (also referred to herein as a "SOC/CRAC nucleic acid") refers to a nucleic acid molecule which: (1) hybridizes under stringent conditions to one or more of the nucleic acids having the sequences of SEQ. ID NOS. 7, 27, 29, and/or 31 (sequences of the mouse and human SOC-2/CRAC-1, human SOC-3/CRAC-2, and human SOC-4/CRAC-3 nucleic acids), and (2) codes for a SOC-2/CRAC-1, a SOC-3/CRAC-2 or a SOC-4/CRAC-3 calcium channel polypeptide, respectively, or unique fragments of said SOC-2/CRAC-1, SOC-3/CRAC-2, or SOC-4/CRAC-3 polypeptide.

As used herein, a SOC/CRAC calcium channel polypeptide (also referred to herein as a "SOC/CRAC polypeptide") refers to a polypeptide that is coded for by a SOC-2/CRAC-1, a SOC-3/CRAC-2, and/or a SOC-4/CRAC-3 nucleic acid. Preferably, the above-identified SOC/CRAC polypeptides mediate transport of calcium into and out of a cell.

SOC/CRAC polypeptides also are useful as immunogenic molecules for the generation of binding polypeptides (e.g., antibodies) which bind selectively to SOC/CRAC (e.g., SOC-2/CRAC-1, SOC-3/CRAC-2, and/or SOC-4/CRAC-3) polypeptides. Such antibodies can be used in diagnostic assays to identify and/or quantify the presence of a SOC/CRAC polypeptide in a sample, such as a biological fluid or biopsy sample. SOC/CRAC polypeptides further embrace functionally equivalent fragments, variants, and analogs of the preferred SOC/CRAC polypeptides, provided that the fragments, variants, and analogs also are useful in mediating calcium transport into and out of intracellular calcium stores.

As used herein, "SOC/CRAC calcium channel activity" refers to Ca^{2+} transport ("Ca²⁺ fluxing") across the plasma membrane that is mediated by a SOC/CRAC calcium channel polypeptide. The SOC/CRAC calcium channel polypeptide typically has one or more of the following properties: high selectivity, a unitary conductance below the detection level of the patch clamp method, and are subject to inhibition by high intracellular calcium levels. Such activity can be easily detected using standard methodology well known in the art. See, e.g., the Examples and Neher, E., "Ion channels for communication between and within cells",

Science, 1992; 256:498-502; and Hoth, M., and Penner, R., "Depletion of intracellular calcium stores activates a calcium current in mast cells", Nature, 1992; 355 (6358):353-6.

According to one aspect of the invention, isolated nucleic acid molecules which code for one or more member(s) of the SOC/CRAC family of calcium channel polypeptides are provided. The isolated nucleic acid molecules are selected from the following groups:

(a) nucleic acid molecules which hybridize under stringent conditions to one or more nucleic acid molecules selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;

(b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;

(c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and

(d) complements of (a), (b) or (c).

In certain embodiments, the isolated nucleic acid molecule comprises one or more of nucleotides 1-1212 of SEQ ID NO:1; nucleotides 1-739 of SEQ ID NO:3; nucleotides 1-1579 of SEQ ID NO:5; nucleotides 1-5117 of SEQ ID NO:23; the mouse homolog for SOC-2/CRAC-1 corresponding to SEQ ID NO:7; nucleotides 1-2180 of SEQ ID NO:25; nucleotides 382-5976 of SEQ ID NO:27; nucleotides 73-3714 of SEQ ID NO:29; and nucleotides 23-3434 of SEQ ID NO:31. In yet other embodiments, the isolated nucleic acid molecule comprises a molecule which encodes a polypeptide having one or more sequences selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32.

According to yet another aspect of the invention, an isolated nucleic acid molecule is provided which is selected from the group consisting of:

(a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, (of sufficient length to represent a sequence unique within the human genome); and (b) complements of (a), provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to a sequence in the prior art as represented by the sequence group consisting of: (1) sequences having the SEQ ID NOs or GenBank accession numbers of Table I, (2) complements of (1), and (3) fragments of (1) and (2).

In some embodiments, the sequence of contiguous nucleotides is selected from the group consisting of (1) at least two contiguous nucleotides nonidentical to the sequence group, (2) at least three contiguous nucleotides nonidentical to the sequence group, (3) at least four contiguous nucleotides nonidentical to the sequence group, (4) at least five contiguous nucleotides nonidentical to the sequence group, (5) at least six contiguous nucleotides nonidentical to the sequence group, (6) at least seven contiguous nucleotides nonidentical to the sequence group.

In other embodiments, the unique fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 40 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, 200 nucleotides, 1000 nucleotides and every integer length therebetween.

According to another aspect of the invention, expression vectors and host cells containing (e.g., transformed or transfected with) expression vectors comprising the nucleic acid molecules disclosed herein operably linked to a promoter are provided. In certain preferred embodiments, the host cells are eukaryotic cells.

The isolated nucleic acid molecules disclosed herein have various utilities, including their use as probes and primers to identify additional members of the SOC/CRAC family of calcium channels, as diagnostic reagents for identifying the presence of SOC/CRAC polypeptides in biological or other samples, and as agents for generating SOC/CRAC binding polypeptides (e.g., antibodies) that can be used as reagents in diagnostic and therapeutic assays to identify the presence, absence, and/or amounts of a SOC/CRAC nucleic acid or polypeptide in a biological or other sample.

As used herein with respect to nucleic acids, the term "isolated" means: (i) amplified *in vitro* by, for example, polymerase chain reaction (PCR); (ii) recombinantly produced by cloning; (iii) purified, as by cleavage and gel separation; or (iv) synthesized by, for example, chemical synthesis. An isolated nucleic acid is one which is readily manipulatable by recombinant DNA techniques well known in the art. Thus, a nucleotide sequence contained in a vector in which 5' and 3' restriction sites are known or for which polymerase chain reaction (PCR) primer sequences have been disclosed is considered isolated but a nucleic acid sequence existing in its native state in its natural host is not. An isolated nucleic acid may be substantially purified, but need not be. For example, a nucleic acid that is isolated within a cloning or expression vector is not pure in that it may comprise only a tiny percentage of the

material in the cell in which it resides. Such a nucleic acid is isolated, however, as the term is used herein because it is readily manipulatable by standard techniques known to those of ordinary skill in the art.

As used herein with respect to polypeptides (discussed below), the term "isolated" means separated from its native environment in sufficiently pure form so that it can be manipulated or used for any one of the purposes of the invention. Thus, isolated means sufficiently pure to be used (i) to raise and/or isolate antibodies, (ii) as a reagent in an assay, or (iii) for sequencing, etc.

Homologs and alleles of the SOC/CRAC nucleic acids of the invention can be identified by conventional techniques. Thus, an aspect of the invention is those nucleic acid sequences which code for SOC/CRAC polypeptides and which hybridize to a nucleic acid molecule selected from a group consisting of the nucleic acid of SEQ ID NO:1, the nucleic acid of SEQ ID NO:3, the nucleic acid of SEQ ID NO:5, the nucleic acid of SEQ ID NO:7, the nucleic acid of SEQ ID NO:23, the nucleic acid of SEQ ID NO:25, the nucleic acid of SEQ ID NO:27, the nucleic acid of SEQ ID NO:29, and the nucleic acid of SEQ ID NO:31, under stringent conditions. The term "stringent conditions" as used herein refers to parameters with which the art is familiar. Nucleic acid hybridization parameters may be found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. More specifically, stringent conditions, as used herein, refers, for example, to hybridization at 65°C in hybridization buffer (3.5 x SSC, 0.02% Ficoll, 0.02% polyvinyl pyrrolidone, 0.02% Bovine Serum Albumin, 2.5mM NaH₂PO₄(pH7), 0.5% SDS, 2mM EDTA). SSC is 0.15M sodium chloride/0.15M sodium citrate, pH7; SDS is sodium dodecyl sulphate; and EDTA is ethylenediaminetetracetic acid. After hybridization, the membrane upon which the DNA is transferred is washed at 2 x SSC at room temperature and then at 0.1 x SSC/0.1 x SDS at temperatures up to 68°C.

There are other conditions, reagents, and so forth which can be used, and would result in a similar degree of stringency. The skilled artisan will be familiar with such conditions, and thus they are not given here. It will be understood, however, that the skilled artisan will be able to manipulate the conditions in a manner to permit the clear identification of homologs and alleles of the SOC/CRAC nucleic acids of the invention. The skilled artisan also is familiar with the methodology for screening cells and libraries for expression of such

molecules which then are routinely isolated, followed by isolation of the pertinent nucleic acid molecule and sequencing.

In general homologs and alleles typically will share at least 40% nucleotide identity and/or at least 50% amino acid identity to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and/or SEQ ID NO:31, and SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, respectively. In some instances sequences will share at least 50% nucleotide identity and/or at least 65% amino acid identity and in still other instances sequences will share at least 60% nucleotide identity and/or at least 75% amino acid identity. The homology can be calculated using various, publicly available software tools developed by NCBI (Bethesda, Maryland) that can be obtained through the internet (<ftp://ncbi.nlm.nih.gov/pub/>). Exemplary tools include the BLAST system available at <http://www.ncbi.nlm.nih.gov>. Pairwise and ClustalW alignments (BLOSUM30 matrix setting) as well as Kyte-Doolittle hydropathic analysis can be obtained using the MacVector sequence analysis software (Oxford Molecular Group). Watson-Crick complements of the foregoing nucleic acids also are embraced by the invention.

In screening for SOC/CRAC related genes, such as homologs and alleles of SOC-2/CRAC-1 and/or SOC-3/CRAC-2, a Southern blot may be performed using the foregoing conditions, together with a radioactive probe. After washing the membrane to which the DNA is finally transferred, the membrane can be placed against X-ray film or a phosphorimager plate to detect the radioactive signal.

Given that the expression of the SOC/CRAC gene is prominent in certain human tissues (e.g., SOC-2/CRAC-1: lymphoid tissue/heart, SOC-3/CRAC-2: kidney/colon, SOC-4/CRAC-3: prostate), and given the teachings herein of partial human SOC/CRAC cDNA clones, full-length and other mammalian sequences corresponding to the human SOC/CRAC partial nucleic acid sequences can be isolated from, for example, a cDNA library prepared from one or more of the tissues in which SOC-2/CRAC-1 expression is prominent, SOC-3/CRAC-2 is prominent, and/or SOC-4/CRAC-3 expression is prominent, using standard colony hybridization techniques.

The invention also includes degenerate nucleic acids which include alternative codons to those present in the native materials. For example, serine residues are encoded by the codons TCA, AGT, TCC, TCG, TCT and AGC. Each of the six codons is equivalent for the purposes of encoding a serine residue. Thus, it will be apparent to one of ordinary skill in the

art that any of the serine-encoding nucleotide triplets may be employed to direct the protein synthesis apparatus, *in vitro* or *in vivo*, to incorporate a serine residue into an elongating SOC/CRAC polypeptide. Similarly, nucleotide sequence triplets which encode other amino acid residues include, but are not limited to: CCA, CCC, CCG and CCT (proline codons); CGA, CGC, CGG, CGT, AGA and AGG (arginine codons); ACA, ACC, ACG and ACT (threonine codons); AAC and AAT (asparagine codons); and ATA, ATC and ATT (isoleucine codons). Other amino acid residues may be encoded similarly by multiple nucleotide sequences. Thus, the invention embraces degenerate nucleic acids that differ from the biologically isolated nucleic acids in codon sequence due to the degeneracy of the genetic code.

The invention also provides isolated unique fragments of an isolated nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31. A unique fragment is one that is a 'signature' for the larger nucleic acid. For example, the unique fragment is long enough to assure that its precise sequence is not found in molecules within the human genome outside of the SOC/CRAC nucleic acids defined above (and human alleles). Those of ordinary skill in the art may apply no more than routine procedures to determine if a fragment is unique within the human genome.

Unique fragments, however, exclude fragments completely composed of the nucleotide sequences of any of GenBank accession numbers and SEQ ID NOs listed in Table I (SEQ ID NO:9, AB001535, AI226731, H18835, AA419592, AA261842, AA419407, AI098310, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853), or other previously published sequences as of the filing date of this application.

A fragment which is completely composed of the sequence described in the foregoing GenBank deposits and SEQ ID NO:9, is one which does not include any of the nucleotides unique to the sequences of the invention. Thus, a unique fragment must contain a nucleotide sequence other than the exact sequence of those in GenBank or fragments thereof. The difference may be an addition, deletion or substitution with respect to the GenBank sequence or it may be a sequence wholly separate from the GenBank sequence.

Unique fragments can be used as probes in Southern and Northern blot assays to identify such nucleic acids, or can be used in amplification assays such as those employing PCR. As known to those skilled in the art, large probes such as 200, 250, 300 or more nucleotides are preferred for certain uses such as Southern and Northern blots, while smaller fragments will be preferred for uses such as PCR. Unique fragments also can be used to produce fusion proteins for generating antibodies or determining binding of the polypeptide fragments, as demonstrated in the Examples, or for generating immunoassay components. Likewise, unique fragments can be employed to produce nonfused fragments of the SOC/CRAC polypeptides, useful, for example, in the preparation of antibodies, immunoassays or therapeutic applications. Unique fragments further can be used as antisense molecules to inhibit the expression of SOC/CRAC nucleic acids and polypeptides, respectively.

As will be recognized by those skilled in the art, the size of the unique fragment will depend upon its conservancy in the genetic code. Thus, some regions of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and complements thereof, will require longer segments to be unique while others will require only short segments, typically between 12 and 32 nucleotides long (e.g. 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31 and 32 bases) or more, up to the entire length of the disclosed sequence. As mentioned above, this disclosure intends to embrace each and every fragment of each sequence, beginning at the first nucleotide, the second nucleotide and so on, up to 8 nucleotides short of the end, and ending anywhere from nucleotide number 8, 9, 10 and so on for each sequence, up to the very last nucleotide, (provided the sequence is unique as described above). Virtually any segment of the region of SEQ ID NO:1 beginning at nucleotide 1 and ending at nucleotide 1212, or SEQ ID NO:3 beginning at nucleotide 1 and ending at nucleotide 739, or SEQ ID NO:5 beginning at nucleotide 1 and ending at nucleotide 1579, or SEQ ID NO:7 beginning at nucleotide 1 and ending at nucleotide 3532, or SEQ ID NO:23 beginning at nucleotide 1 and ending at nucleotide 5117, SEQ ID NO:25 beginning at nucleotide 1 and ending at nucleotide 2180, SEQ ID NO:27 beginning at nucleotide 1 and ending at nucleotide 7419, or SEQ ID NO:29 beginning at nucleotide 1 and ending at nucleotide 4061, or SEQ ID NO:31 beginning at nucleotide 1 and ending at nucleotide 4646, or complements thereof, that is 20 or more nucleotides in length will be unique. Those skilled in the art are well versed in methods for selecting such sequences, typically on the basis of the ability of the unique

fragment to selectively distinguish the sequence of interest from other sequences in the human genome of the fragment to those on known databases typically is all that is necessary, although *in vitro* confirmatory hybridization and sequencing analysis may be performed.

As mentioned above, the invention embraces antisense oligonucleotides that selectively bind to a nucleic acid molecule encoding a SOC/CRAC polypeptide, to decrease SOC/CRAC calcium channel activity. When using antisense preparations of the invention, slow intravenous administration is preferred.

As used herein, the term "antisense oligonucleotide" or "antisense" describes an oligonucleotide that is an oligoribonucleotide, oligodeoxyribonucleotide, modified oligoribonucleotide, or modified oligodeoxyribonucleotide which hybridizes under physiological conditions to DNA comprising a particular gene or to an mRNA transcript of that gene and, thereby, inhibits the transcription of that gene and/or the translation of that mRNA. The antisense molecules are designed so as to interfere with transcription or translation of a target gene upon hybridization with the target gene or transcript. Those skilled in the art will recognize that the exact length of the antisense oligonucleotide and its degree of complementarity with its target will depend upon the specific target selected, including the sequence of the target and the particular bases which comprise that sequence. It is preferred that the antisense oligonucleotide be constructed and arranged so as to bind selectively with the target under physiological conditions, i.e., to hybridize substantially more to the target sequence than to any other sequence in the target cell under physiological conditions. Based upon SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, or upon allelic or homologous genomic and/or cDNA sequences, one of skill in the art can easily choose and synthesize any of a number of appropriate antisense molecules for use in accordance with the present invention. In order to be sufficiently selective and potent for inhibition, such antisense oligonucleotides should comprise at least 10 and, more preferably, at least 15 consecutive bases which are complementary to the target, although in certain cases modified oligonucleotides as short as 7 bases in length have been used successfully as antisense oligonucleotides (Wagner et al., *Nat. Med.* 1(11):1116-1118, 1995). Most preferably, the antisense oligonucleotides comprise a complementary sequence of 20-30 bases. Although oligonucleotides may be chosen which are antisense to any region of the gene or mRNA transcripts, in preferred embodiments the antisense oligonucleotides correspond to N-terminal or 5' upstream sites such as translation initiation, transcription initiation or promoter sites. In

addition, 3'-untranslated regions may be targeted by antisense oligonucleotides. Targeting to mRNA splicing sites has also been used in the art but may be less preferred if alternative mRNA splicing occurs. In addition, the antisense is targeted, preferably, to sites in which mRNA secondary structure is not expected (see, e.g., Sainio et al., *Cell Mol. Neurobiol.* 14(5):439-457, 1994) and at which proteins are not expected to bind. Finally, although, SEQ ID No:1 discloses a cDNA sequence, one of ordinary skill in the art may easily derive the genomic DNA corresponding to this sequence. Thus, the present invention also provides for antisense oligonucleotides which are complementary to the genomic DNA corresponding to SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31. Similarly, antisense to allelic or homologous SOC/CRAC cDNAs and genomic DNAs are enabled without undue experimentation.

In one set of embodiments, the antisense oligonucleotides of the invention may be composed of "natural" deoxyribonucleotides, ribonucleotides, or any combination thereof. That is, the 5' end of one native nucleotide and the 3' end of another native nucleotide may be covalently linked, as in natural systems, via a phosphodiester internucleoside linkage. These oligonucleotides may be prepared by art recognized methods which may be carried out manually or by an automated synthesizer. They also may be produced recombinantly by vectors.

In preferred embodiments, however, the antisense oligonucleotides of the invention also may include "modified" oligonucleotides. That is, the oligonucleotides may be modified in a number of ways which do not prevent them from hybridizing to their target but which enhance their stability or targeting or which otherwise enhance their therapeutic effectiveness.

The term "modified oligonucleotide" as used herein describes an oligonucleotide in which (1) at least two of its nucleotides are covalently linked via a synthetic internucleoside linkage (i.e., a linkage other than a phosphodiester linkage between the 5' end of one nucleotide and the 3' end of another nucleotide) and/or (2) a chemical group not normally associated with nucleic acids has been covalently attached to the oligonucleotide. Preferred synthetic internucleoside linkages are phosphorothioates, alkylphosphonates, phosphorodithioates, phosphate esters, alkylphosphonothioates, phosphoramidates, carbamates, carbonates, phosphate triesters, acetamides, carboxymethyl esters and peptides.

The term "modified oligonucleotide" also encompasses oligonucleotides with a covalently modified base and/or sugar. For example, modified oligonucleotides include

-20-

oligonucleotides having backbone sugars which are covalently attached to low molecular weight organic groups other than a hydroxyl group at the 3' position and other than a phosphate group at the 5' position. Thus modified oligonucleotides may include a 2'-O-alkylated ribose group. In addition, modified oligonucleotides may include sugars such as arabinose instead of ribose. The present invention, thus, contemplates pharmaceutical preparations containing modified antisense molecules that are complementary to and hybridizable with, under physiological conditions, nucleic acids encoding SOC/CRAC polypeptides, together with pharmaceutically acceptable carriers. Antisense oligonucleotides may be administered as part of a pharmaceutical composition. Such a pharmaceutical composition may include the antisense oligonucleotides in combination with any standard physiologically and/or pharmaceutically acceptable carriers which are known in the art. The compositions should be sterile and contain a therapeutically effective amount of the antisense oligonucleotides in a unit of weight or volume suitable for administration to a patient. The term "pharmaceutically acceptable" means a non-toxic material that does not interfere with the effectiveness of the biological activity of the active ingredients. The term "physiologically acceptable" refers to a non-toxic material that is compatible with a biological system such as a cell, cell culture, tissue, or organism. The characteristics of the carrier will depend on the route of administration. Physiologically and pharmaceutically acceptable carriers include diluents, fillers, salts, buffers, stabilizers, solubilizers, and other materials which are well known in the art.

The invention also involves expression vectors coding for SOC/CRAC proteins and fragments and variants thereof and host cells containing those expression vectors. Virtually any cells, prokaryotic or eukaryotic, which can be transformed with heterologous DNA or RNA and which can be grown or maintained in culture, may be used in the practice of the invention. Examples include bacterial cells such as *E.coli* and eukaryotic cells such as mouse, hamster, pig, goat, primate, yeast, xenopous, etc. They may be of a wide variety of tissue types, including mast cells, fibroblasts, oocytes and lymphocytes, and they may be primary cells or cell lines. Specific examples include CHO cells and COS cells. Cell-free transcription systems also may be used in lieu of cells.

As used herein, a "vector" may be any of a number of nucleic acids into which a desired sequence may be inserted by restriction and ligation for transport between different genetic environments or for expression in a host cell. Vectors are typically composed of DNA although RNA vectors are also available. Vectors include, but are not limited to,

plasmids, phagemids and virus genomes. A cloning vector is one which is able to replicate in a host cell, and which is further characterized by one or more endonuclease restriction sites at which the vector may be cut in a determinable fashion and into which a desired DNA sequence may be ligated such that the new recombinant vector retains its ability to replicate in the host cell. In the case of plasmids, replication of the desired sequence may occur many times as the plasmid increases in copy number within the host bacterium or just a single time per host before the host reproduces by mitosis. In the case of phage, replication may occur actively during a lytic phase or passively during a lysogenic phase. An expression vector is one into which a desired DNA sequence may be inserted by restriction and ligation such that it is operably joined to regulatory sequences and may be expressed as an RNA transcript. Vectors may further contain one or more marker sequences suitable for use in the identification of cells which have or have not been transformed or transfected with the vector. Markers include, for example, genes encoding proteins which increase or decrease either resistance or sensitivity to antibiotics or other compounds, genes which encode enzymes whose activities are detectable by standard assays known in the art (e.g., β -galactosidase or alkaline phosphatase), and genes which visibly affect the phenotype of transformed or transfected cells, hosts, colonies or plaques (e.g., green fluorescent protein). Preferred vectors are those capable of autonomous replication and expression of the structural gene products present in the DNA segments to which they are operably joined.

As used herein, a coding sequence and regulatory sequences are said to be "operably" joined when they are covalently linked in such a way as to place the expression or transcription of the coding sequence under the influence or control of the regulatory sequences. If it is desired that the coding sequences be translated into a functional protein, two DNA sequences are said to be operably joined if induction of a promoter in the 5' regulatory sequences results in the transcription of the coding sequence and if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region to direct the transcription of the coding sequences, or (3) interfere with the ability of the corresponding RNA transcript to be translated into a protein. Thus, a promoter region would be operably joined to a coding sequence if the promoter region were capable of effecting transcription of that DNA sequence such that the resulting transcript might be translated into the desired protein or polypeptide.

The precise nature of the regulatory sequences needed for gene expression may vary between species or cell types, but shall in general include, as necessary, 5' non-transcribed

and 5' non-translated sequences involved with the initiation of transcription and translation respectively, such as a TATA box, capping sequence, CAAT sequence, and the like. Especially, such 5' non-transcribed regulatory sequences will include a promoter region which includes a promoter sequence for transcriptional control of the operably joined gene.

5 Regulatory sequences may also include enhancer sequences or upstream activator sequences as desired. The vectors of the invention may optionally include 5' leader or signal sequences. The choice and design of an appropriate vector is within the ability and discretion of one of ordinary skill in the art.

According to yet another aspect of the invention, isolated SOC/CRAC polypeptides

10 are provided. Preferably, the isolated SOC/CRAC polypeptides are encoded by the isolated SOC/CRAC nucleic acid molecules disclosed herein. More preferably, the isolated SOC/CRAC polypeptides of the invention are encoded by the nucleic acid molecules having SEQ ID Nos. 1, 3, 5, 7, 23, 25, 27, 29, and 31. In yet other embodiments, the isolated SOC/CRAC polypeptides of the invention have an amino acid sequence selected from the

15 group consisting of SEQ ID Nos. 2, 4, 6, 8, 24, 26, 28, 30 and 32. Preferably, the isolated SOC/CRAC polypeptides are of sufficient length to represent a sequence unique within the human genome. Thus, the preferred embodiments include a sequence of contiguous amino acids which is not identical to a prior art sequence as represented by the sequence group consisting of the contiguous amino acids identified in Table II (SEQ ID NO:10, SEQ ID

20 NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19 and GenBank Acc. Nos. AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572).

In certain embodiments, the isolated SOC/CRAC polypeptides are immunogenic and

25 can be used to generate binding polypeptides (e.g., antibodies) for use in diagnostic and therapeutic applications. Such binding polypeptides also are useful for detecting the presence, absence, and/or amounts of a SOC/CRAC nucleic acid or polypeptide in a sample such as a biological fluid or biopsy sample. Preferably, the SOC/CRAC polypeptides that are useful for generating binding polypeptides are unique polypeptides and, therefore, binding of the

30 antibody to a SOC/CRAC polypeptide in a sample is selective for the SOC/CRAC polypeptide.

Expression vectors containing all the necessary elements for expression are commercially available and known to those skilled in the art. See, e.g., Sambrook et al.,

Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, 1989. Cells are genetically engineered by the introduction into the cells of heterologous DNA (RNA) encoding a SOC/CRAC polypeptide or fragment or variant thereof. The heterologous DNA (RNA) is placed under operable control of transcriptional elements to permit the expression of the heterologous DNA in the host cell.

Preferred systems for mRNA expression in mammalian cells are those such as pRc/CMV (available from Invitrogen, Carlsbad, CA) that contain a selectable marker such as a gene that confers G418 resistance (which facilitates the selection of stably transfected cell lines) and the human cytomegalovirus (CMV) enhancer-promoter sequences. Additionally, suitable for expression in primate or canine cell lines is the pCEP4 vector (Invitrogen, Carlsbad, CA), which contains an Epstein Barr virus (EBV) origin of replication, facilitating the maintenance of plasmid as a multicopy extrachromosomal element. Another expression vector is the pEF-BOS plasmid containing the promoter of polypeptide Elongation Factor 1 α , which stimulates efficiently transcription *in vitro*. The plasmid is described by Mishizuma and Nagata (*Nuc. Acids Res.* 18:5322, 1990), and its use in transfection experiments is disclosed by, for example, Demoulin (*Mol. Cell. Biol.* 16:4710-4716, 1996). Still another preferred expression vector is an adenovirus, described by Stratford-Perricaudet, which is defective for E1 and E3 proteins (*J. Clin. Invest.* 90:626-630, 1992). The use of the adenovirus as an Adeno.P1A recombinant is disclosed by Warnier et al., in intradermal injection in mice for immunization against P1A (*Int. J. Cancer*, 67:303-310, 1996).

The invention also embraces so-called expression kits, which allow the artisan to prepare a desired expression vector or vectors. Such expression kits include at least separate portions of each of the previously discussed coding sequences. Other components may be added, as desired, as long as the previously mentioned sequences, which are required, are included.

It will also be recognized that the invention embraces the use of the above described, SOC/CRAC cDNA sequence containing expression vectors, to transfect host cells and cell lines, by these prokaryotic (e.g., *E. coli*), or eukaryotic (e.g., CHO cells, COS cells, yeast expression systems and recombinant baculovirus expression in insect cells). Especially useful are mammalian cells such as mouse, hamster, pig, goat, primate, etc. They may be of a wide variety of tissue types, and include primary cells and cell lines. Specific examples include dendritic cells, U293 cells, peripheral blood leukocytes, bone marrow stem cells and embryonic stem cells. The invention also permits the construction of SOC/CRAC gene

“knock-outs” in cells and in animals, providing materials for studying certain aspects of SOC/CRAC calcium channel activity.

The invention also provides isolated polypeptides (including whole proteins and partial proteins), encoded by the foregoing SOC/CRAC nucleic acids, and include the polypeptides of SEQ ID NO:2, 4, 6, 8, 24, 26, 28, 30, 32, and unique fragments thereof. Such polypeptides are useful, for example, to regulate calcium transport-mediated cell growth, differentiation and proliferation, to generate antibodies, as components of immunoassays, etc. Polypeptides can be isolated from biological samples including tissue or cell homogenates, and can also be expressed recombinantly in a variety of prokaryotic and eukaryotic expression systems by constructing an expression vector appropriate to the expression system, introducing the expression vector into the expression system, and isolating the recombinantly expressed protein. Short polypeptides, including antigenic peptides (such as are presented by MHC molecules on the surface of a cell for immune recognition) also can be synthesized chemically using well-established methods of peptide synthesis.

A unique fragment of a SOC/CRAC polypeptide, in general, has the features and characteristics of unique fragments as discussed above in connection with nucleic acids. As will be recognized by those skilled in the art, the size of the unique fragment will depend upon factors such as whether the fragment constitutes a portion of a conserved protein domain. Thus, some regions of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, will require longer segments to be unique while others will require only short segments, typically between 5 and 12 amino acids (e.g. 5, 6, 7, 8, 9, 10, 11 and 12 amino acids long or more, including each integer up to the full length, >1,000 amino acids long). Virtually any segment of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32, excluding the ones that share identity with it (the polypeptides identified in Table II - SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, and GenBank Acc. Nos. AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572) that is 9 or more amino acids in length will be unique.

Unique fragments of a polypeptide preferably are those fragments which retain a distinct functional capability of the polypeptide. Functional capabilities which can be retained in a unique fragment of a polypeptide include Ca^{2+} fluxing, high selectivity, a unitary

conductance below the detection level of the patch clamp method, and/or and are subject to inhibition by high intracellular calcium levels.

One important aspect of a unique fragment is its ability to act as a signature for identifying the polypeptide. Optionally, another aspect of a unique fragment is its ability to provide an immune response in an animal. Those skilled in the art are well versed in methods for selecting unique amino acid sequences, typically on the basis of the ability of the unique fragment to selectively distinguish the sequence of interest from non-family members. A comparison of the sequence of the fragment to those on known databases typically is all that is necessary.

The invention embraces variants of the SOC/CRAC polypeptides described above. As used herein, a "variant" of a SOC/CRAC polypeptide is a polypeptide which contains one or more modifications to the primary amino acid sequence of a SOC/CRAC polypeptide. Modifications which create a SOC/CRAC polypeptide variant are typically made to the nucleic acid which encodes the SOC/CRAC polypeptide, and can include deletions, point mutations, truncations, amino acid substitutions and addition of amino acids or non-amino acid moieties to: 1) reduce or eliminate a calcium channel activity of a SOC/CRAC polypeptide; 2) enhance a property of a SOC/CRAC polypeptide, such as protein stability in an expression system or the stability of protein-protein binding; 3) provide a novel activity or property to a SOC/CRAC polypeptide, such as addition of an antigenic epitope or addition of a detectable moiety; or 4) to provide equivalent or better binding to a SOC/CRAC polypeptide receptor or other molecule. Alternatively, modifications can be made directly to the polypeptide, such as by cleavage, addition of a linker molecule, addition of a detectable moiety, such as biotin, addition of a fatty acid, and the like. Modifications also embrace fusion proteins comprising all or part of the SOC/CRAC amino acid sequence. One of skill in the art will be familiar with methods for predicting the effect on protein conformation of a change in protein sequence, and can thus "design" a variant SOC/CRAC polypeptide according to known methods. One example of such a method is described by Dahiyat and Mayo in *Science* 278:82-87, 1997, whereby proteins can be designed *de novo*. The method can be applied to a known protein to vary only a portion of the polypeptide sequence. By applying the computational methods of Dahiyat and Mayo, specific variants of a SOC/CRAC calcium channel polypeptide can be proposed and tested to determine whether the variant retains a desired conformation.

5 Variants can include SOC/CRAC polypeptides which are modified specifically to alter a feature of the polypeptide unrelated to its physiological activity. For example, cysteine residues can be substituted or deleted to prevent unwanted disulfide linkages. Similarly, certain amino acids can be changed to enhance expression of a SOC/CRAC polypeptide by eliminating proteolysis by proteases in an expression system (e.g., dibasic amino acid residues in yeast expression systems in which KEX2 protease activity is present).

10 Mutations of a nucleic acid which encodes a SOC/CRAC polypeptide preferably preserve the amino acid reading frame of the coding sequence and, preferably, do not create regions in the nucleic acid which are likely to hybridize to form secondary structures, such as hairpins or loops, which can be deleterious to expression of the variant polypeptide.

15 Mutations can be made by selecting an amino acid substitution, or by random mutagenesis of a selected site in a nucleic acid which encodes the polypeptide. Variant polypeptides are then expressed and tested for one or more activities to determine which mutation provides a variant polypeptide with the desired properties. Further mutations can be made to variants (or to non-variant SOC/CRAC polypeptides) which are silent as to the amino acid sequence of the polypeptide, but which provide preferred codons for translation in a particular host. The preferred codons for translation of a nucleic acid in, e.g., *E. coli*, are well known to those of ordinary skill in the art. Still other mutations can be made to the noncoding sequences of a SOC/CRAC gene or cDNA clone to enhance expression of the polypeptide.

20 The skilled artisan will realize that conservative amino acid substitutions may be made in SOC/CRAC polypeptides to provide functionally equivalent variants of the foregoing polypeptides, i.e., the variants retain the functional capabilities of the SOC/CRAC polypeptides. As used herein, a "conservative amino acid substitution" refers to an amino acid substitution which does not alter the relative charge or size characteristics of the protein in which the amino acid substitution is made. Variants can be prepared according to methods for altering polypeptide sequence known to one of ordinary skill in the art such as are found in references which compile such methods, e.g. *Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York. Exemplary functionally equivalent variants of the SOC/CRAC polypeptides include conservative amino acid substitutions of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and/or SEQ ID NO:32. Conservative substitutions of amino acids

30

include substitutions made amongst amino acids within the following groups: (a) M, I, L, V; (b) F, Y, W; (c) K, R, H; (d) A, G; (e) S, T; (f) Q, N; and (g) E, D.

Thus functionally equivalent variants of SOC/CRAC polypeptides, i.e., variants of SOC/CRAC polypeptides which retain the function of the natural SOC/CRAC polypeptides, are contemplated by the invention. Conservative amino-acid substitutions in the amino acid sequence of SOC/CRAC polypeptides to produce functionally equivalent variants of SOC/CRAC polypeptides typically are made by alteration of a nucleic acid encoding SOC/CRAC polypeptides (e.g., SEQ ID NOs:1, 3, 5, 7, 23, 25, 27, 29, 31). Such substitutions can be made by a variety of methods known to one of ordinary skill in the art. For example, amino acid substitutions may be made by PCR-directed mutation, site-directed mutagenesis according to the method of Kunkel (Kunkel, *Proc. Nat. Acad. Sci. U.S.A.* 82: 488-492, 1985), or by chemical synthesis of a gene encoding a SOC/CRAC polypeptide. The activity of functionally equivalent fragments of SOC/CRAC polypeptides can be tested by cloning the gene encoding the altered SOC/CRAC polypeptide into a bacterial or mammalian expression vector, introducing the vector into an appropriate host cell, expressing the altered SOC/CRAC polypeptide, and testing for a functional capability of the SOC/CRAC polypeptides as disclosed herein (e.g., SOC/CRAC calcium channel activity).

The invention as described herein has a number of uses, some of which are described elsewhere herein. First, the invention permits isolation of SOC/CRAC polypeptides, including the isolation of the complete SOC/CRAC polypeptide. A variety of methodologies well-known to the skilled practitioner can be utilized to obtain isolated SOC/CRAC molecules. The polypeptide may be purified from cells which naturally produce the polypeptide by chromatographic means or immunological recognition. Alternatively, an expression vector may be introduced into cells to cause production of the polypeptide. In another method, mRNA transcripts may be microinjected or otherwise introduced into cells to cause production of the encoded polypeptide. Translation of SOC/CRAC mRNA in cell-free extracts such as the reticulocyte lysate system also may be used to produce SOC/CRAC polypeptides. Those skilled in the art also can readily follow known methods for isolating SOC/CRAC polypeptides. These include, but are not limited to, immunochromatography, HPLC, size-exclusion chromatography, ion-exchange chromatography and immune-affinity chromatography.

The invention also provides, in certain embodiments, "dominant negative" polypeptides derived from SOC/CRAC polypeptides. A dominant negative polypeptide is an

inactive variant of a protein, which, by interacting with the cellular machinery, displaces an active protein from its interaction with the cellular machinery or competes with the active protein, thereby reducing the effect of the active protein. For example, a dominant negative receptor which binds a ligand but does not transmit a signal in response to binding of the ligand can reduce the biological effect of expression of the ligand. Likewise, a dominant negative inactive SOC/CRAC calcium channel which interacts normally with the cell membrane but which does not mediate calcium transport can reduce calcium transport in a cell. Similarly, a dominant negative transcription factor which binds to a promoter site in the control region of a gene but does not increase gene transcription can reduce the effect of a normal transcription factor by occupying promoter binding sites without increasing transcription.

The end result of the expression of a dominant negative polypeptide in a cell is a reduction in function of active proteins. One of ordinary skill in the art can assess the potential for a dominant negative variant of a protein, and using standard mutagenesis techniques to create one or more dominant negative variant polypeptides. See, e.g., U.S. Patent No. 5,580,723 and Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, Second Edition, Cold Spring Harbor Laboratory Press. The skilled artisan then can test the population of mutagenized polypeptides for diminution in a selected and/or for retention of such an activity. Other similar methods for creating and testing dominant negative variants of a protein will be apparent to one of ordinary skill in the art.

According to another aspect, the invention provides a method for isolating a SOC/CRAC molecule having SOC/CRAC calcium channel activity. The method involves contacting a binding molecule that is a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing one or more SOC/CRAC molecules under conditions that allow such binding (see earlier discussion) to form a complex, detecting the presence of the complex, isolating the SOC/CRAC molecule from the complex, and determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity. Thus, the invention is useful for identifying and isolating full length complementary (cDNA) or genomic nucleic acids encoding SOC/CRAC polypeptides having SOC/CRAC calcium channel activity. Identification and isolation of such nucleic acids and polypeptides may be accomplished by hybridizing/binding, under appropriate conditions well known in the art, libraries and/or restriction enzyme-digested human nucleic acids, with a labeled SOC/CRAC molecular probe. As used herein, a "label" includes molecules that are incorporated into, for

example, a SOC/CRAC molecule (nucleic acid or peptide), that can be directly or indirectly detected. A wide variety of detectable labels are well known in the art that can be used, and include labels that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc), or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseshoe peroxidase, etc.). The label may be bound to a SOC/CRAC binding partner, or incorporated into the structure of the binding partner.

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly detected through optical or electron density, radioactive emissions, nonradioactive energy transfers, etc. or indirectly detected with antibody conjugates, streptavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art. Once a library clone or hybridizing fragment is identified in the hybridization/binding reaction, it can be further isolated by employing standard isolation/cloning techniques known to those of skill in the art. See, generally, Sambrook et al., 1989, *Molecular Cloning: A Laboratory Manual*, 2nd Edition, Cold Spring Harbor Laboratory Press. In addition, nucleic acid amplification techniques well known in the art, may also be used to locate splice variants of calcium channel (or calcium channel subunits) with SOC/CRAC calcium channel activity. Size and sequence determinations of the amplification products can reveal splice variants.

The foregoing isolated nucleic acids and polypeptides may then be compared to the nucleic acids and polypeptides of the present invention in order to identify homogeneity or divergence of the sequences, and be further characterized functionally to determine whether they belong to a family of molecules with SOC/CRAC calcium channel activity (for methodology see under the Examples section).

The isolation of the SOC/CRAC cDNA and/or partial sequences thereof also makes it possible for the artisan to diagnose a disorder characterized by an aberrant expression of SOC/CRAC. These methods involve determining expression of the SOC/CRAC gene, and/or SOC/CRAC polypeptides derived therefrom. In the former situation, such determinations can be carried out via any standard nucleic acid determination assay, including the polymerase chain reaction, or assaying with labeled hybridization probes as exemplified below. In the latter situation, such determination can be carried out via any standard immunological assay using, for example, antibodies which bind to the SOC/CRAC protein.

The invention also embraces isolated peptide binding agents which, for example, can be antibodies or fragments of antibodies ("binding polypeptides"), having the ability to selectively bind to SOC/CRAC polypeptides. Antibodies include polyclonal and monoclonal antibodies, prepared according to conventional methodology. In certain embodiments, the invention excludes binding agents (e.g., antibodies) that bind to the polypeptides encoded by the nucleic acids of SEQ ID NOs: 10, 12, 13, 14, 15, 17, and 19.

Significantly, as is well-known in the art, only a small portion of an antibody molecule, the paratope, is involved in the binding of the antibody to its epitope (see, in general, Clark, W.R. (1986) The Experimental Foundations of Modern Immunology Wiley & Sons, Inc., New York; Roitt, I. (1991) Essential Immunology, 7th Ed., Blackwell Scientific Publications, Oxford). The pFc' and Fc regions, for example, are effectors of the complement cascade but are not involved in antigen binding. An antibody from which the pFc' region has been enzymatically cleaved, or which has been produced without the pFc' region, designated an F(ab')₂ fragment, retains both of the antigen binding sites of an intact antibody. Similarly, an antibody from which the Fc region has been enzymatically cleaved, or which has been produced without the Fc region, designated an Fab fragment, retains one of the antigen binding sites of an intact antibody molecule. Proceeding further, Fab fragments consist of a covalently bound antibody light chain and a portion of the antibody heavy chain denoted Fd. The Fd fragments are the major determinant of antibody specificity (a single Fd fragment may be associated with up to ten different light chains without altering antibody specificity) and Fd fragments retain epitope-binding ability in isolation.

Within the antigen-binding portion of an antibody, as is well-known in the art, there are complementarity determining regions (CDRs), which directly interact with the epitope of the antigen, and framework regions (FRs), which maintain the tertiary structure of the paratope (see, in general, Clark, 1986; Roitt, 1991). In both the heavy chain Fd fragment and the light chain of IgG immunoglobulins, there are four framework regions (FR1 through FR4) separated respectively by three complementarity determining regions (CDR1 through CDR3). The CDRs, and in particular the CDR3 regions, and more particularly the heavy chain CDR3, are largely responsible for antibody specificity.

It is now well-established in the art that the non-CDR regions of a mammalian antibody may be replaced with similar regions of conspecific or heterospecific antibodies while retaining the epitopic specificity of the original antibody. This is most clearly manifested in the development and use of "humanized" antibodies in which non-human CDRs

are covalently joined to human FR and/or Fc/pFc' regions to produce a functional antibody. Thus, for example, PCT International Publication Number WO 92/04381 teaches the production and use of humanized murine RSV antibodies in which at least a portion of the murine FR regions have been replaced by FR regions of human origin. Such antibodies, including fragments of intact antibodies with antigen-binding ability, are often referred to as "chimeric" antibodies.

Thus, as will be apparent to one of ordinary skill in the art, the present invention also provides for F(ab')₂, Fab, Fv and Fd fragments; chimeric antibodies in which the Fc and/or FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric F(ab')₂ fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; chimeric Fab fragment antibodies in which the FR and/or CDR1 and/or CDR2 and/or light chain CDR3 regions have been replaced by homologous human or non-human sequences; and chimeric Fd fragment antibodies in which the FR and/or CDR1 and/or CDR2 regions have been replaced by homologous human or non-human sequences. The present invention also includes so-called single chain antibodies.

Thus, the invention involves binding polypeptides of numerous size and type that bind selectively to SOC/CRAC polypeptides, and complexes containing SOC/CRAC polypeptides. These binding polypeptides also may be derived also from sources other than antibody technology. For example, such polypeptide binding agents can be provided by degenerate peptide libraries which can be readily prepared in solution, in immobilized form, as bacterial flagella peptide display libraries or as phage display libraries. Combinatorial libraries also can be synthesized of peptides containing one or more amino acids. Libraries further can be synthesized of peptides and non-peptide synthetic moieties.

Phage display can be particularly effective in identifying binding peptides useful according to the invention. Briefly, one prepares a phage library (using e.g. m13, fd, or lambda phage), displaying inserts from 4 to about 80 amino acid residues using conventional procedures. The inserts may represent, for example, a completely degenerate or biased array. One then can select phage-bearing inserts which bind to the SOC/CRAC polypeptide or a complex containing a SOC/CRAC polypeptide. This process can be repeated through several cycles of reselection of phage that bind to the SOC/CRAC polypeptide or complex. Repeated rounds lead to enrichment of phage bearing particular sequences. DNA sequence analysis can be conducted to identify the sequences of the expressed polypeptides. The minimal linear

portion of the sequence that binds to the SOC/CRAC polypeptide or complex can be determined. One can repeat the procedure using a biased library containing inserts containing part or all of the minimal linear portion plus one or more additional degenerate residues upstream or downstream thereof. Yeast two-hybrid screening methods also may be used to
5 identify polypeptides that bind to the SOC/CRAC polypeptides. Thus, the SOC/CRAC polypeptides of the invention, or a fragment thereof, or complexes of SOC/CRAC can be used to screen peptide libraries, including phage display libraries, to identify and select peptide binding polypeptides that selectively bind to the SOC/CRAC polypeptides of the invention. Such molecules can be used, as described, for screening assays, for purification protocols, for
10 interfering directly with the functioning of SOC/CRAC and for other purposes that will be apparent to those of ordinary skill in the art.

A SOC/CRAC polypeptide, or a fragment thereof, also can be used to isolate naturally occurring, polypeptide binding partners which may associate with the SOC/CRAC polypeptide in the membrane of a cell. Isolation of binding partners may be performed
15 according to well-known methods. For example, isolated SOC/CRAC polypeptides can be attached to a substrate, and then a solution suspected of containing an SOC/CRAC binding partner may be applied to the substrate. If the binding partner for SOC/CRAC polypeptides is present in the solution, then it will bind to the substrate-bound SOC/CRAC polypeptide. The binding partner then may be isolated. Other proteins which are binding partners for
20 SOC/CRAC, may be isolated by similar methods without undue experimentation.

The invention also provides novel kits which could be used to measure the levels of the nucleic acids of the invention, expression products of the invention or anti-SOC/CRAC antibodies. In the case of nucleic acid detection, pairs of primers for amplifying SOC/CRAC nucleic acids can be included. The preferred kits would include controls such as known
25 amounts of nucleic acid probes, SOC/CRAC epitopes (such as SOC/CRAC expression products) or anti-SOC/CRAC antibodies, as well as instructions or other printed material. In certain embodiments the printed material can characterize risk of developing a disorder that is characterized by aberrant SOC/CRAC polypeptide expression based upon the outcome of the assay. The reagents may be packaged in containers and/or coated on wells in predetermined
30 amounts, and the kits may include standard materials such as labeled immunological reagents (such as labeled anti-IgG antibodies) and the like. One kit is a packaged polystyrene microtiter plate coated with a SOC/CRAC polypeptide and a container containing labeled anti-human IgG antibodies. A well of the plate is contacted with, for example, serum, washed

and then contacted with the anti-IgG antibody. The label is then detected. A kit embodying features of the present invention is comprised of the following major elements: packaging an agent of the invention, a control agent, and instructions. Packaging is a box-like structure for holding a vial (or number of vials) containing an agent of the invention. a vial (or number of vials) containing a control agent, and instructions. Individuals skilled in the art can readily modify packaging to suit individual needs.

Another aspect of the invention is a method for determining the level of SOC/CRAC expression in a subject. As used herein, a subject is a human, non-human primate, cow, horse, pig, sheep, goat, dog, cat or rodent. In all embodiments, human subjects are preferred. Expression is defined either as SOC/CRAC mRNA expression or SOC/CRAC polypeptide expression. Various methods can be used to measure expression. Preferred embodiments of the invention include PCR and Northern blotting for measuring mRNA expression, and monoclonal or polyclonal SOC/CRAC antisera as reagents to measure SOC/CRAC polypeptide expression. In certain embodiments, test samples such as biopsy samples, and biological fluids such as blood, are used as test samples. SOC/CRAC expression in a test sample of a subject is compared to SOC/CRAC expression in control sample to, e.g., assess the presence or absence or stage of a proliferative disorder (e.g., a lymphocyte proliferative disorder) in a subject.

SOC/CRAC polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts. Recombinantly produced SOC/CRAC polypeptides include chimeric proteins comprising a fusion of a SOC/CRAC protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the SOC/CRAC polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein. A polypeptide fused to a SOC/CRAC polypeptide or fragment may also provide means of readily detecting the fusion protein, e.g., by immunological recognition or by fluorescent labeling.

The invention is also useful in the generation of transgenic non-human animals. As used herein, "transgenic non-human animals" includes non-human animals having one or more exogenous nucleic acid molecules incorporated in germ line cells and/or somatic cells. Thus the transgenic animal include "knockout" animals having a homozygous or heterozygous gene disruption by homologous recombination, animals having episomal or chromosomally incorporated expression vectors, etc. Knockout animals can be prepared by

homologous recombination using embryonic stem cells as is well known in the art. The recombination may be facilitated using, for example, the cre/lox system or other recombinase systems known to one of ordinary skill in the art. In certain embodiments, the recombinase system itself is expressed conditionally, for example, in certain tissues or cell types, at certain embryonic or post-embryonic developmental stages, inducibly by the addition of a compound which increases or decreases expression, and the like. In general, the conditional expression vectors used in such systems use a variety of promoters which confer the desired gene expression pattern (e.g., temporal or spatial). Conditional promoters also can be operably linked to SOC/CRAC nucleic acid molecules to increase expression of SOC/CRAC in a regulated or conditional manner. *Trans*-acting negative regulators of SOC/CRAC calcium channel activity or expression also can be operably linked to a conditional promoter as described above. Such *trans*-acting regulators include antisense SOC/CRAC nucleic acids molecules, nucleic acid molecules which encode dominant negative SOC/CRAC molecules, ribozyme molecules specific for SOC/CRAC nucleic acids, and the like. The transgenic non-human animals are useful in experiments directed toward testing biochemical or physiological effects of diagnostics or therapeutics for conditions characterized by increased or decreased SOC/CRAC expression. Other uses will be apparent to one of ordinary skill in the art.

The invention further provides efficient methods of identifying agents or lead compounds for agents active at the level of a SOC/CRAC polypeptide (e.g., a SOC/CRAC polypeptide) or SOC/CRAC fragment dependent cellular function. In particular, such functions include interaction with other polypeptides or fragments thereof, and selective binding to certain molecules (e.g., agonists and antagonists). Generally, the screening methods involve assaying for compounds which interfere with SOC/CRAC calcium channel activity, although compounds which enhance SOC/CRAC calcium channel activity also can be assayed using the screening methods. Such methods are adaptable to automated, high throughput screening of compounds. The target therapeutic indications for pharmacological agents detected by the screening methods are limited only in that the target cellular function be subject to modulation by alteration of the formation of a complex comprising a SOC/CRAC polypeptide or fragment thereof and one or more SOC/CRAC binding targets. Target indications include cellular processes modulated by SOC/CRAC such as Ca^{2+} fluxing, and affected by SOC/CRAC ability to form complexes with other molecules and polypeptides as, for example, may be present in the cell membrane.

A wide variety of assays for pharmacological agents are provided, including, expression assays, labeled *in vitro* protein-protein binding assays, electrophoretic mobility shift assays, immunoassays, cell-based assays such as calcium transport assays, etc. For example, two-hybrid screens are used to rapidly examine the effect of transfected nucleic acids on the intracellular binding of SOC/CRAC or SOC/CRAC fragments to specific intracellular targets (e.g. a tyrosine kinase). The transfected nucleic acids can encode, for example, combinatorial peptide libraries or cDNA libraries. Convenient reagents for such assays, e.g., GAL4 fusion proteins, are known in the art. An exemplary cell-based assay involves transfecting a cell with a nucleic acid encoding a SOC/CRAC polypeptide fused to a GAL4 DNA binding domain and a nucleic acid encoding a reporter gene operably linked to a gene expression regulatory region, such as one or more GAL4 binding sites. Activation of reporter gene transcription occurs when the SOC/CRAC and reporter fusion polypeptides bind such as to enable transcription of the reporter gene. Agents which modulate a SOC/CRAC polypeptide mediated cell function are then detected through a change in the expression of reporter gene. Methods for determining changes in the expression of a reporter gene are known in the art.

In an expression system, for example, a SOC/CRAC polypeptide is attached to a membrane, the membrane preferably separating two fluid environments and being otherwise not permeable to Ca^{2+} . Such separation is preferred so that a change in Ca^{2+} concentration on either side of the membrane is mediated only through the attached SOC/CRAC polypeptide. Preferably, a SOC/CRAC polypeptide is expressed in an intact cell and is present on the cell-membrane (as in physiologic conditions). The cell expressing the SOC/CRAC polypeptide is preferably a eukaryotic cell, and the SOC/CRAC polypeptide is preferably recombinantly expressed, although cells naturally expressing a SOC/CRAC polypeptide may also be used. Synthetic membranes, however, containing SOC/CRAC polypeptides may also be used. See, e.g., K. Kiselyov, et al., Functional interaction between InsP3 receptors and store-operated Htrp3 channels, Nature 396, 478-82 (1998).

The cell expressing the SOC/CRAC polypeptide is incubated under conditions which, in the absence of the candidate agent, permit calcium flux into the cell and allow detection of a reference calcium concentration. For example, depletion of intracellular calcium stores with thapsigargin or other agents (Putney, J.W. Jr., in Capacitative Calcium Entry, R.G. Landes Co. and Chapman & Hall, 1997) would produce a given level of SOC/CRAC channel activation and a given reference calcium concentration. Detection of a decrease in the

foregoing activities (i.e., a decrease in the intracellular calcium concentration) relative to the reference calcium concentration indicates that the candidate agent is a lead compound for an agent to inhibit SOC/CRAC calcium channel activity. Preferred SOC/CRAC polypeptides include the polypeptides of claim 15.

5 SOC/CRAC fragments used in the methods, when not produced by a transfected nucleic acid are added to an assay mixture as an isolated polypeptide. SOC/CRAC polypeptides preferably are produced recombinantly, although such polypeptides may be isolated from biological extracts or chemically synthesized. Recombinantly produced SOC/CRAC polypeptides include chimeric proteins comprising a fusion of a SOC/CRAC
10 protein with another polypeptide, e.g., a polypeptide capable of providing or enhancing protein-protein binding, sequence specific nucleic acid binding (such as GAL4), enhancing stability of the SOC/CRAC polypeptide under assay conditions, or providing a detectable moiety, such as green fluorescent protein or Flag epitope.

The assay mixture is comprised of a SOC/CRAC polypeptide binding target
15 (candidate agent) capable of interacting with a SOC/CRAC polypeptide. While natural SOC/CRAC binding targets may be used, it is frequently preferred to use portions (e.g., peptides or nucleic acid fragments) or analogs (i.e., agents which mimic the SOC/CRAC binding properties of the natural binding target for purposes of the assay) of the SOC/CRAC binding target so long as the portion or analog provides binding affinity and avidity to the
20 SOC/CRAC polypeptide (or fragment thereof) measurable in the assay.

The assay mixture also comprises a candidate agent (binding target, e.g., agonist/antagonist). Typically, a plurality of assay mixtures are run in parallel with different agent concentrations to obtain a different response to the various concentrations. Typically, one of these concentrations serves as a negative control, i.e., at zero concentration of agent or
25 at a concentration of agent below the limits of assay detection. Candidate agents encompass numerous chemical classes, although typically they are organic compounds. Preferably, the candidate agents are small organic compounds, i.e., those having a molecular weight of more than 50 yet less than about 2500, preferably less than about 1000 and, more preferably, less than about 500. Candidate agents comprise functional chemical groups necessary for
30 structural interactions with polypeptides and/or nucleic acids, and typically include at least an amine, carbonyl, hydroxyl or carboxyl group, preferably at least two of the functional chemical groups and more preferably at least three of the functional chemical groups. The candidate agents can comprise cyclic carbon or heterocyclic structure and/or aromatic or

polyaromatic structures substituted with one or more of the above-identified functional groups. Candidate agents also can be biomolecules such as peptides, saccharides, fatty acids, sterols, isoprenoids, purines, pyrimidines, derivatives or structural analogs of the above, or combinations thereof and the like. Where the agent is a nucleic acid, the agent typically is a DNA or RNA molecule, although modified nucleic acids as defined herein are also contemplated.

Candidate agents are obtained from a wide variety of sources including libraries of synthetic or natural compounds. For example, numerous means are available for random and directed synthesis of a wide variety of organic compounds and biomolecules, including expression of randomized oligonucleotides, synthetic organic combinatorial libraries, phage display libraries of random peptides, and the like. Alternatively, libraries of natural compounds in the form of bacterial, fungal, plant and animal extracts are available or readily produced. Additionally, natural and synthetically produced libraries and compounds can be readily modified through conventional chemical, physical, and biochemical means. Further, known agents may be subjected to directed or random chemical modifications such as acylation, alkylation, esterification, amidification, etc. to produce structural analogs of the agents. Non-SOC/CRAC calcium channel agonists and antagonists, for example, include agents such as dihydropyridines (DHPs), phenylalkylamines, omega conotoxin (omega-CgTx) and pyrazonoylguanidines.

A variety of other reagents also can be included in the mixture. These include reagents such as salts, buffers, neutral proteins (e.g., albumin), detergents, etc. which may be used to facilitate optimal protein-protein, protein-nucleic acid, and/or protein/membrane component binding association. Such a reagent may also reduce non-specific or background interactions of the reaction components. Other reagents that improve the efficiency of the assay such as protease, inhibitors, nuclease inhibitors, antimicrobial agents, and the like may also be used.

The mixture of the foregoing assay materials is incubated under conditions whereby, but for the presence of the candidate agent, the SOC/CRAC polypeptide specifically binds the cellular binding target, a portion thereof or analog thereof. The order of addition of components, incubation temperature, time of incubation, and other perimeters of the assay may be readily determined. Such experimentation merely involves optimization of the assay parameters, not the fundamental composition of the assay. Incubation temperatures typically

are between 4°C and 40°C. Incubation times preferably are minimized to facilitate rapid, high throughput screening, and typically are between 0.1 and 10 hours.

After incubation, the presence or absence of specific binding between the SOC/CRAC polypeptide and one or more binding targets is detected by any convenient method available to the user. For cell free binding type assays, a separation step is often used to separate bound from unbound components. The separation step may be accomplished in a variety of ways. Conveniently, at least one of the components is immobilized on a solid substrate, from which the unbound components may be easily separated. The solid substrate can be made of a wide variety of materials and in a wide variety of shapes, e.g., microtiter plate, microbead, dipstick, resin particle, etc. The substrate preferably is chosen to maximum signal to noise ratios, primarily to minimize background binding, as well as for ease of separation and cost.

Separation may be effected for example, by removing a bead or dipstick from a reservoir, emptying or diluting a reservoir such as a microtiter plate well, rinsing a bead, particle, chromatographic column or filter with a wash solution or solvent. The separation step preferably includes multiple rinses or washes. For example, when the solid substrate is a microtiter plate, the wells may be washed several times with a washing solution, which typically includes those components of the incubation mixture that do not participate in specific bindings such as salts, buffer, detergent, non-specific protein, etc. Where the solid substrate is a magnetic bead, the beads may be washed one or more times with a washing solution and isolated using a magnet.

Detection may be effected in any convenient way for cell-based assays such as two- or three-hybrid screens. The transcript resulting from a reporter gene transcription assay of SOC/CRAC polypeptide interacting with a target molecule typically encodes a directly or indirectly detectable product, e.g., β -galactosidase activity, luciferase activity, and the like. For cell-free binding assays, one of the components usually comprises, or is coupled to, a detectable label. A wide variety of labels can be used, such as those that provide direct detection (e.g., radioactivity, luminescence, optical or electron density, etc.) or indirect detection (e.g., epitope tag such as the FLAG epitope, enzyme tag such as horseradish peroxidase, etc.). The label may be bound to a SOC/CRAC binding partner, or incorporated into the structure of the binding partner.

A variety of methods may be used to detect the label, depending on the nature of the label and other assay components. For example, the label may be detected while bound to the solid substrate or subsequent to separation from the solid substrate. Labels may be directly

detected through optical or electron density, radioactive emissions, nonradiative energy transfers, etc. or indirectly detected with antibody conjugates, strepavidin-biotin conjugates, etc. Methods for detecting the labels are well known in the art.

Of particular importance in any of the foregoing assays and binding studies is the use of a specific sequence motif identified in the SOC-2/CRAC-1 polypeptide sequence as a kinase catalytic domain. According to the invention, amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24) (or a fragment thereof), show a localized homology with the catalytic domains of eukaryotic elongation factor-2 kinase (eEF-2 kinase, GenBank Acc. no. U93850) and *Dictyostelium* myocin heavy chain kinase A (MHCK A, GenBank Acc. no. U16856), as disclosed in Ryazanov AG, et al., *Proc Natl Acad Sci U S A*, 1997, 94(10):4884-4889. Therefore, according to the invention, a method for identifying agents useful in the modulation of SOC/CRAC polypeptide kinase activity is provided. The method involves contacting a SOC/CRAC polypeptide with kinase activity, that includes, for example, amino acids 999-1180 of the SOC-2/CRAC-1 polypeptide (SEQ ID NO:24) with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity; detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and comparing the kinase activity in the previous step with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates (increases or decreases) SOC/CRAC kinase activity. Other controls for kinase activity can also be performed at the same time, for example, by utilizing eEF-2 kinase and/or *Dictyostelium* MHC Kinase A, in a similar manner to the SOC/CRAC member. Methods for performing such kinase activity assays are well known in the art.

The invention thus provides SOC/CRAC-specific binding agents, methods of identifying and making such agents, and their use in diagnosis, therapy and pharmaceutical development. For example, SOC/CRAC-specific agents are useful in a variety of diagnostic and therapeutic applications, especially where disease or disease prognosis is associated with altered SOC/CRAC and SOC/CRAC calcium channel fluxing characteristics. Novel SOC/CRAC-specific binding agents include SOC/CRAC-specific antibodies and other natural intracellular and extracellular binding agents identified with assays such as two hybrid screens, and non-natural intracellular and extracellular binding agents identified in screens of chemical libraries and the like.

In general, the specificity of SOC/CRAC binding to a specific molecule is determined by binding equilibrium constants. Targets which are capable of selectively binding a SOC/CRAC polypeptide preferably have binding equilibrium constants of at least about 10^7 M^{-1} , more preferably at least about 10^8 M^{-1} , and most preferably at least about 10^9 M^{-1} . The wide variety of cell based and cell free assays may be used to demonstrate SOC/CRAC-specific binding. Cell based assays include one, two and three hybrid screens, assays in which SOC/CRAC-mediated transcription is inhibited or increased, etc. Cell free assays include SOC/CRAC-protein binding assays, immunoassays, etc. Other assays useful for screening agents which bind SOC/CRAC polypeptides include fluorescence resonance energy transfer (FRET), and electrophoretic mobility shift analysis (EMSA).

Various techniques may be employed for introducing nucleic acids of the invention into cells, depending on whether the nucleic acids are introduced *in vitro* or *in vivo* in a host. Such techniques include transfection of nucleic acid- $CaPO_4$ precipitates, transfection of nucleic acids associated with DEAE, transfection with a retrovirus including the nucleic acid of interest, liposome mediated transfection, and the like. For certain uses, it is preferred to target the nucleic acid to particular cells. In such instances, a vehicle used for delivering a nucleic acid of the invention into a cell (e.g., a retrovirus, or other virus; a liposome) can have a targeting molecule attached thereto. For example, a molecule such as an antibody specific for a surface membrane protein on the target cell or a ligand for a receptor on the target cell can be bound to or incorporated within the nucleic acid delivery vehicle. For example, where liposomes are employed to deliver the nucleic acids of the invention, proteins which bind to a surface membrane protein associated with endocytosis may be incorporated into the liposome formulation for targeting and/or to facilitate uptake. Such proteins include capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half life, and the like. Polymeric delivery systems also have been used successfully to deliver nucleic acids into cells, as is known by those skilled in the art. Such systems even permit oral delivery of nucleic acids.

Other delivery systems can include time-release, delayed release or sustained release delivery systems. Such systems can avoid repeated administrations of the anti-inflammatory agent, increasing convenience to the subject and the physician. Many types of release delivery systems are available and known to those of ordinary skill in the art. They include polymer base systems such as poly(lactide-glycolide), copolyoxalates, polycaprolactones,

polyesteramides, polyorthoesters, polyhydroxybutyric acid, and polyanhydrides. Microcapsules of the foregoing polymers containing drugs are described in, for example, U.S. Patent 5,075,109. Delivery systems also include non-polymer systems that are: lipids including sterols such as cholesterol, cholesterol esters and fatty acids or neutral fats such as mono- di- and tri-glycerides; hydrogel release systems; systatic systems; peptide based systems; wax coatings; compressed tablets using conventional binders and excipients; partially fused implants; and the like. Specific examples include, but are not limited to: (a) erosional systems in which an agent of the invention is contained in a form within a matrix such as those described in U.S. Patent Nos. 4,452,775, 4,675,189, and 5,736,152, and (b) diffusional systems in which an active component permeates at a controlled rate from a polymer such as described in U.S. Patent Nos. 3,854,480, 5,133,974 and 5,407,686. In addition, pump-based hardware delivery systems can be used, some of which are adapted for implantation.

Use of a long-term sustained release implant may be particularly suitable for treatment of chronic conditions. Long-term release, as used herein, means that the implant is constructed and arranged to deliver therapeutic levels of the active ingredient for at least 30 days, and preferably 60 days. Long-term sustained release implants are well-known to those of ordinary skill in the art and include some of the release systems described above.

The invention also contemplates gene therapy. The procedure for performing *ex vivo* gene therapy is outlined in U.S. Patent 5,399,346 and in exhibits submitted in the file history of that patent, all of which are publicly available documents. In general, it involves introduction *in vitro* of a functional copy of a gene into a cell(s) of a subject which contains a defective copy of the gene, and returning the genetically engineered cell(s) to the subject. The functional copy of the gene is under operable control of regulatory elements which permit expression of the gene in the genetically engineered cell(s). Numerous transfection and transduction techniques as well as appropriate expression vectors are well known to those of ordinary skill in the art, some of which are described in PCT application WO95/00654. *In vivo* gene therapy using vectors such as adenovirus, retroviruses, herpes virus, and targeted liposomes also is contemplated according to the invention. See, e.g., U.S. Patent Nos. 5,670,488, entitled "Adenovirus Vector for Gene Therapy", issued to Gregory et al., and 5,672,344, entitled "Viral-Mediated Gene Transfer System", issued to Kelley et al.

The invention will be more fully understood by reference to the following examples. These examples, however, are merely intended to illustrate the embodiments of the invention and are not to be construed to limit the scope of the invention.

Examples

5 As an initial approach to identifying SOC/CRAC channels, we considered publicly available data and hypothesized that the following characteristics are likely to be exhibited by SOC/CRAC calcium channels: i) SOC/CRAC calcium channels would be integral membrane proteins related (probably distantly) to one of the known calcium channel families (e.g. voltage gated, ligand gated, Trp), and therefore should have a pore region formed by a tetramer of 6-7 transmembrane (TM) regions; ii) high calcium selectivity was likely to come at the price of complexity, and therefore these were likely to be large proteins; iii) the high calcium selectivity of this type of channel was likely to be useful and, therefore, highly conserved; and iv) these channels should be expressed in one or more types of lymphocytes, since ICRAC is best defined in those cell types. Since the full genome of the nematode *C. elegans* is nearing completion, and IP3-dependent calcium signals have recently been shown to be required for one or more aspects of *C. elegans* development, we took the set of proteins encoded by this genome (at the time this search was initiated WORMPEP14 was the available predicted protein set) and began searching for proteins which fit the criteria above. This search began by proceeding in alphabetical order through WORMPEP14 and arbitrarily excluding all proteins below approximately 1000 amino acids in size, followed by focusing on remaining proteins with clear TM spanning regions similar to those of other calcium channels. We stopped this screen on encountering a protein designated C05C12.3, a predicted protein of 1816 amino acids (SEQ ID NO:13). C05C12.3 was notable because its central pore region had some sequence similarity to but was clearly distinct from members of the Trp family of calcium channels, and the hydrophobicity plot of this region showed a characteristically wide spacing between the fifth and sixth TM regions for the amino acid residues which are thought to line the channel pore region and mediate the calcium selectivity of the channels. In addition, it lacked any ankyrin repeats in the region amino-terminal to its pore region, further distinguishing it from other Trp family proteins.

30 We then used C05C12.3 for BLAST alignment screening of the rest of the *C. elegans* genome and also mammalian databases for homologous proteins, revealing two other *C. elegans* homologues (SEQ ID NO:14 and SEQ ID NO:15), and also a recently cloned mammalian protein named melastatin-1 (MLSN-1/SOC-1, SEQ ID NOs:9 and 10, and

GenBank Acc. No. AF071787). Using these sequences, we subsequently performed an exhaustive screening of publicly accessible EST databases in search of lymphocyte homologues, but were unsuccessful in detecting any homologous transcripts in any lymphocyte lines. Since MLSN-1 (SEQ ID NOs:9 and 10) was expressed exclusively in melanocytes and retina by Northern blot hybridization and by EST database searching, there was no evidence that this type of channel was expressed in the type of cell in which ICRAC-like currents were best defined. Subsequent BLAST searches picked up mouse EST sequence AI098310 (SEQ ID NO:22) from a monocyte cell line. The I.M.A.G.E. consortium clone containing the above-identified EST was then purchased from ATCC (clone ID. 1312756, Manassas, VA) and was further characterized. Using other portions of this sequence in EST searches, we subsequently picked up similar sequences in human B-cells (SEQ ID NOs:20 and 21), and other cell types as well (SEQ ID NOs: 11, 12, 16, 17, 18, and 19). Most of these sequences were subsequently identified to be part of the 3'-UTR or of the carboxy terminal region of the proteins, which are not readily identifiable as Trp channels, providing an explanation for the art's inability to detect any type of Trp related transcripts in lymphocytes. Partial sequences from the 5' and/or 3' ends of the above identified clones were then used to screen leukocyte and kidney cDNA libraries to extend the original sequences more toward the 5' and/or 3' ends.

In view of the foregoing, it was concluded that channels of this type were expressed in many types of lymphocytes, and therefore were members of a new family of SOC/CRAC calcium channels.

Experimental Procedures

Screening of the cDNA libraries

Leukocyte and kidney cDNA libraries from Life Technologies (Gaithersburg, MD) were screened using the Gene Trapper II methodology (Life Technologies) according to manufacturer's recommendation, using the inserts of I.M.A.G.E. clone ID nos. 1312756 and 1076485 from ATCC (Manassas, VA), under stringent hybridization conditions. Using standard methodology (*Molecular Cloning: A Laboratory Manual*, J. Sambrook, et al., eds., Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1989, or *Current Protocols in Molecular Biology*, F.M. Ausubel, et al., eds., John Wiley & Sons, Inc., New York), individual cDNA clones were subjected to 3-4 rounds of amplification and purification under the same hybridization conditions.

After excision from the vector and subcloning of inserts into the plasmid forms, several clones were sequenced by the Beth Israel Deaconess Medical Center's Automated

Sequencing Facility. Molecular biological techniques such as restriction enzyme treatment, subcloning, DNA extraction, bacterial culture and purification of DNA fragments were performed according to methods well known in the art. Computer analyses of protein and DNA sequences was done using "Assemblylign" (Oxford Molecular, Cambell, CA). Multiple
5 alignments of the SOC/CRAC family members were produced using the CLUSTAL facility of the MacVector program. Restriction endonucleases, expression vectors, and modifying enzymes were purchased from commercial sources (Gibco-BRL). Sequencing vectors for DNA were purchased from Stratagene (La Jolla, CA).

Once the first members of what appeared to be a novel family of calcium channel
10 receptors were identified and characterized, additional BLAST alignments were performed with the newly characterized nucleic acid sequences. An initial match was with genomic DNA fragment NH0332L11 (Genbank Acc. No. AC005538). Using this genomic sequence, promoters were designed and a number of cDNA libraries was surveyed by PCR. A prostate specific message was identified and characterized, leading to the isolation and
15 characterization of SOC-4/CRAC-3 (SEQ ID NOs: 31 and 32).

Functional Assays

Transient Expression of SOC/CRAC

In our initial transient expression experiments, we expressed or expect to express a SOC/CRAC molecule transiently in RBL-2H3 mast cells, Jurkat T cells, and A20
20 B-lymphocytes using both electroporation and vaccinia virus-driven expression, and measured the calcium influx produced by depletion of intracellular calcium stores with thapsigargin. Each of the foregoing techniques is well known to those of ordinary skill in the art and can be performed using various methods (see, e.g., Current Methods in Molecular Biology, eds. Ausubal, F.M., et al. 1987, Green Publishers and Wiley Interscience, N.Y.,
25 N.Y.). Exemplary methods are described herein.

Depletion of intracellular calcium stores is accomplished by treating the cells with 1 micromolar thapsigargin; alternative agents which function to deplete intracellular stores are described in by Putney, J.W. Jr., in Capacitative Calcium Entry, R.G. Landes Co. and Chapman & Hall, 1997 and include, for example, ionomycin, cyclopiazonic acid, and DBHQ.

30 Calcium influx is determined by measuring cytoplasmic calcium as indicated using the fura-2 fluorescent calcium indicator (see, e.g., G. Grynkiewicz, M. Poenie, R. Y. Tsien, A new generation of Ca²⁺ indicators with greatly improved fluorescence properties, J. Biol

Chem 260, 3440-50 (1985), and M. Poenie, R. Tsien, Fura-2: a powerful new tool for measuring and imaging $[Ca^{2+}]_i$ in single cells, Prog Clin Biol Res 210, 53-6 (1986)).

Patch Clamp Analysis and Determining Selectivity of SOC/CRAC

Patch clamp analysis of cells injected with SOC/CRAC cRNA is performed by using the general patch technique as described in Neher, E., "Ion channels for communication between and within cells", Science, 1992; 256:498-502. Specific techniques for applying the patch clamp analysis to RBL cells are described in Hoth, M., and Penner, R., "Depletion of intracellular calcium stores activates a calcium current in mast cells", Nature, 1992; 355:3535-355. Additional protocols for applying the patch clamp technique to other cell types are described in Putney, J.W. Jr., in Capacitative Calcium Entry, R.G. Landes Co. and Chapman & Hall, 1997

An exemplary protocol for patch clamp analysis of SOC/CRAC molecule expressed in RBL-2H3 mast cells using a recombinant vaccinia virus is as follows. The currents elicited by store depletion are determined using the whole cell configuration (Neher, E., Science, 1992; 256:498-502). Currents in SOC/CRAC expressing cells are compared to currents in control cells expressing an irrelevant protein or a classic Trp family calcium channel known as VR1 (M. J. Caterina, et al., The capsaicin receptor: a heat-activated ion channel in the pain pathway [see comments], Nature 389, 816-24 (1997)) in order to assess the contribution of SOC/CRAC expression. In addition, the magnitude of whole cell currents in the presence of extracellular calcium (10 mM), barium (10 mM), or magnesium (10 mM) are compared to determine the relative permeability of the channels to each of these ions (Hoth, M., and Penner, R., Nature, 1992; 355:3535-355) and, thereby, determine the ionic selectivity.

Pharmacologic Behavior of SOC/CRAC

For analysis of the pharmacologic behavior of a SOC/CRAC molecule, a SOC/CRAC molecule is expressed in RBL-2H3 mast cells using a recombinant vaccinia virus, and the degree of calcium influx elicited by store depletion is monitored using a bulk spectrofluorimeter or a fluorescence microscope and the calcium sensitive dye fura-2 (G. Grynkiewicz, M. Poenie, R. Y. Tsien, A new generation of Ca^{2+} indicators with greatly improved fluorescence properties, J Biol Chem 260, 3440-50 (1985) and M. Poenie, R. Tsien, Fura-2: a powerful new tool for measuring and imaging $[Ca^{2+}]_i$ in single cells, Prog Clin Biol Res 210, 53-6 (1986)). The level of cytoplasmic calcium in SOC/CRAC expressing cells is compared to the level achieved in control cells expressing an irrelevant protein or a classic Trp. family calcium channels known as VR1 (M. J. Caterina, et al., The

capsaicin receptor: a heat-activated ion channel in the pain pathway [see comments], Nature 389, 816-24 (1997)). These cells then are pre-incubated with the desired pharmacologic reagent, and again the response to store depletion is monitored. Comparison of the effect of depleting stores in SOC/CRAC expressing cells relative to controls in the presence or absence of the pharmacologic reagent is used to assess the ability of that reagent to modulate SOC/CRAC activity. Sphingosine is an exemplary molecule that can be used as pharmacologic reagents for pharmacologic characterization of SOC/CRAC calcium channels. See, e.g., Mathes, C., et al., Calcium release activated calcium current as a direct target for sphingosine, J Biol Chem 273(39):25020-25030 (1998). Other non-specific calcium channel inhibitors that can be used for this purpose include SKR96365 (Calbiochem) and Lanthanum.

Bulk Calcium Assays

Bulk calcium assays can be performed in a PTI Deltascan bulk spectrofluorometer using fura-2 as described in Scharenberg AM, et al., EMBO J, 1995, 14(14):3385-94.

Gene Targeting

The method (and reagents) described by Buerstedde JM et al, (Cell, 1991, Oct 4;67(1):179-88), was used to generate "knockouts" in cells. Briefly, part of the chicken SOC-2/CRAC-1 genomic sequence coding for the transmembrane region was cloned utilizing the human sequence as the probe in a chicken library screen. Chicken SOC-2/CRAC-1 clones were isolated and characterized using standard methodology. The putative exon and domain arrangement of the chicken SOC-2/CRAC-1, is depicted in Figure 1. The exons coding for TM5 (pore region) and TM6, were replaced with promoter/antibiotic cassettes (see Figure1). These targeting vectors were then used to target (and replace) the endogenous gene in DT-40 cells (chicken B lymphocyte cells).

Results

Example 1: Transient Expression of SOC/CRAC

In the above-identified cell lines and using both of the foregoing expression techniques, SOC/CRAC expression enhances thapsigargin-dependent influx. In addition, SOC/CRAC expression also enhances the amount of intracellular calcium stores. That this effect is likely due to SOC/CRAC acting as a plasma membrane calcium channel can be confirmed by producing an in-frame carboxy-terminal translational fusion with green fluorescent protein followed by confocal microscopy, revealing that SOC/CRAC is expressed predominantly as a plasma membrane calcium channel.

Example 2: Patch Clamp Analysis

The biophysical characteristics of SOC/CRAC enhanced currents when expressed in *Xenopus* oocytes are determined. SOC/CRAC cRNA injection is able to enhance thapsigargin-dependent whole cell currents. In addition, SOC/CRAC does not alter the reversal potential of these currents and the determination of the P_{Ca}/P_{Na} ratio shows that SOC/CRAC channels are highly calcium selective.

Example 3: *Pharmacologic Behavior of SOC/CRAC*

The pharmacologic behavior of SOC/CRAC is evaluated as described above. SOC/CRAC-enhanced influx is inhibited by sphingosine in a manner that is substantially the same as that of endogenous thapsigargin-dependent calcium influx.

Example 4: *Gene targeting*

Transfection of DT-40 cells with the foregoing targeting vectors, selection for antibiotic resistance, and screening, is collectively referred to, herein, as a round of targeting. For the first round of targeting SOC-2/CRAC-1, 18/24 clones with homologous recombination of the targeting construct into one of the endogenous SOC-2/CRAC-1 alleles were obtained. On the second round of targeting (in order to target the second allele and therefore generate a homozygous SOC-2/CRAC-1 mutant cell), 0/48 clones were obtained. These results indicate that a "null" SOC-2/CRAC-1 mutation is detrimental to DT-40 cells, and that SOC-2/CRAC-1 is required for cell viability.

Table I. Nucleotide Sequences with homologies to SOC/CRAC nucleic acids

Sequences with SEQ ID NOs and GenBank accession numbers:
SEQ ID NO:9, AB001535, AI226731, H18835, AA419592, AA261842, AA419407, AA592910, D86107, AI098310, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, AC005538, AA654650, AA370110, AA313170, AA493512, AI670079, AI671853.

Table II. Amino Acid Sequences with homologies to SOC/CRAC polypeptides

Sequences with SEQ ID NOs and GenBank accession numbers:
SEQ ID NO:10, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:17, SEQ ID NO:19, AB001535, AA592910, D86107, AF071787, Z77132, Z83117, Z68333, AA708532, AA551759, AA932133, R47363, N31660, NP003298, CAB00861, NP002411, CAA92726, CAB05572.

All references, patents, and patent documents disclosed herein are incorporated by reference herein in their entirety.

What is claimed is presented below and is followed by a Sequence Listing. We claim:

-48-
Claims

1. An isolated nucleic acid molecule, comprising:

(a) nucleic acid molecules which hybridize under stringent conditions to a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31, and which code for a SOC/CRAC polypeptide;

(b) deletions, additions and substitutions of (a) which code for a respective SOC/CRAC polypeptide;

(c) nucleic acid molecules that differ from the nucleic acid molecules of (a) or (b) in codon sequence due to the degeneracy of the genetic code, and

(d) complements of (a), (b) or (c).

2. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:1.

3. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:27.

4. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:29.

5. The isolated nucleic acid molecule of claim 1, wherein the isolated nucleic acid molecule comprises SEQ ID NO:31.

6. An isolated nucleic acid molecule selected from the group consisting of

(a) a unique fragment of a nucleic acid molecule selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:29, and SEQ ID NO:31,

(b) complements of (a),

provided that the unique fragment includes a sequence of contiguous nucleotides which is not identical to any sequence selected from a sequence group consisting of

(1) sequences having the SEQ. ID NOS. or GenBank accession numbers of Table I,

(2) complements of (1), and

(3) fragments of (1) and (2).

7. The isolated nucleic acid molecule of claim 6, wherein the sequence of contiguous nucleotides is selected from the group consisting of:

- (1) at least two contiguous nucleotides nonidentical to the sequence group,
- (2) at least three contiguous nucleotides nonidentical to the sequence group,
- (3) at least four contiguous nucleotides nonidentical to the sequence group,
- (4) at least five contiguous nucleotides nonidentical to the sequence group,
- (5) at least six contiguous nucleotides nonidentical to the sequence group,
- (6) at least seven contiguous nucleotides nonidentical to the sequence group.

8. The isolated nucleic acid molecule of claim 6, wherein the unique fragment has a size selected from the group consisting of at least: 8 nucleotides, 10 nucleotides, 12 nucleotides, 14 nucleotides, 16 nucleotides, 18 nucleotides, 20, nucleotides, 22 nucleotides, 24 nucleotides, 26 nucleotides, 28 nucleotides, 30 nucleotides, 50 nucleotides, 75 nucleotides, 100 nucleotides, and 200 nucleotides.

9. The isolated nucleic acid molecule of claim 6, wherein the molecule encodes a polypeptide which is immunogenic.

10. An expression vector comprising the isolated nucleic acid molecule of claims 1, 2, 3, 4, 5, 6, 7, 8, or 9 operably linked to a promoter.

11. A host cell transformed or transfected with the expression vector of claim 10.

12. An isolated polypeptide encoded by the isolated nucleic acid molecule according to anyone of claims 1 or 6, wherein the polypeptide comprises a SOC/CRAC polypeptide or a unique fragment thereof.

13. The isolated polypeptide of claim 12, wherein the isolated polypeptide is encoded by the isolated nucleic acid molecule of claim 2, 3, 4, or 5.

14. The isolated polypeptide of claim 13, wherein the isolated polypeptide comprises a polypeptide having the sequence of amino acids selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32.

15. An isolated polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3, 4, or 5, wherein the polypeptide, or unique fragment thereof is immunogenic.

16. An isolated binding polypeptide which binds selectively to a polypeptide encoded by the isolated nucleic acid molecule of claim 1, 2, 3, 4, or 5.

5 17. The isolated binding polypeptide of claim 16, wherein the isolated binding polypeptide binds to a polypeptide having the sequence of amino acids selected from the group consisting of SEQ ID NO:2, SEQ ID NO:4, SEQ ID NO:6, SEQ ID NO:8, SEQ ID NO:24, SEQ ID NO:26, SEQ ID NO:28, SEQ ID NO:30, and SEQ ID NO:32.

10 18. The isolated binding polypeptide of claim 17, wherein the isolated binding polypeptide is an antibody or an antibody fragment selected from the group consisting of a Fab fragment, a F(ab)₂ fragment or a fragment including a CDR3 region selective for the polypeptide.

15 19. An isolated polypeptide, comprising a unique fragment of the polypeptide of claim 12 of sufficient length to represent a sequence unique within the human genome, provided that the fragment excludes a sequence of contiguous amino acids identified in Table II.

20. A method for isolating a SOC/CRAC molecule having SOC/CRAC calcium channel activity, comprising:

20 a) contacting a binding molecule that is a SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing one or more SOC/CRAC molecules, under conditions sufficient to form a complex of the SOC/CRAC nucleic acid or the SOC/CRAC binding polypeptide and the SOC/CRAC molecule;

b) detecting the presence of the complex;

c) isolating the SOC/CRAC molecule from the complex; and

25 d) determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity.

21. The method of claim 20, wherein the binding molecule is a SOC/CRAC nucleic acid.

22. The method of claim 20, wherein the binding molecule is a SOC/CRAC binding polypeptide.

23. The method of claim 21, wherein the SOC/CRAC nucleic acid comprises at least 14 nucleotides from any contiguous portion of a sequence of nucleotides selected from the group consisting of SEQ ID NO:1, SEQ ID NO:3, SEQ ID NO:5, SEQ ID NO:7, SEQ ID NO:23, SEQ ID NO:25, SEQ ID NO:27, SEQ ID NO:29, and SEQ ID NO:31.

5 24. A method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity, comprising:

a) contacting a SOC/CRAC polypeptide with a candidate agent suspected of modulating SOC/CRAC calcium channel activity, under conditions sufficient to allow the SOC/CRAC polypeptide to interact selectively with the candidate agent;

10 b) detecting a Ca^{2+} concentration associated with SOC/CRAC calcium channel activity of the SOC/CRAC polypeptide in the presence of the candidate agent; and

c) comparing the Ca^{2+} concentration of step (b) with a control Ca^{2+} concentration of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates SOC/CRAC calcium channel activity.

15 25. A method for determining the level of SOC/CRAC expression in a subject, comprising:

a) measuring the expression of SOC/CRAC in a test sample obtained from the subject, and

20 b) comparing the measured expression of SOC/CRAC in the test sample to the expression of the SOC/CRAC polypeptide in a control to determine the level of SOC/CRAC expression in the subject.

25 26. The method of claim 25, wherein the expression of SOC/CRAC in (b) is SOC/CRAC mRNA expression.

27. The method of claim 25, wherein the expression of SOC/CRAC in (b) is SOC/CRAC polypeptide expression.

28. The method of claim 25, wherein the test sample is tissue.

29. The method of claim 25, wherein the test sample is a biological fluid.

30. The method of claim 26, wherein SOC/CRAC mRNA expression is measured using the Polymerase Chain Reaction (PCR).

31. The method of claim 26, wherein SOC/CRAC mRNA expression is measured using a method selected from the group consisting of northern blotting, monoclonal antisera to SOC/CRAC and polyclonal antisera to SOC/CRAC.

32. A kit, comprising a package containing:

an agent that selectively binds to the isolated nucleic acid of claim 1 or an expression product thereof, and

a control for comparing to a measured value of binding of said agent to said isolated nucleic acid of claim 1 or expression product thereof.

33. The kit of claim 32, wherein the control comprises an epitope of the expression product of the nucleic acid of claim 1.

34. A pharmaceutical composition comprising:

a pharmaceutically effective amount of an agent comprising of an isolated nucleic acid molecule of claim 1 or an expression product thereof, and

a pharmaceutically acceptable carrier.

35. The pharmaceutical composition of claim 34, wherein the agent is an expression product of the isolated nucleic acid molecule of claim 1.

36. A method for identifying agents useful in the modulation of a SOC/CRAC polypeptide kinase activity, comprising:

a) contacting a SOC/CRAC polypeptide with kinase activity with a candidate agent suspected of modulating SOC/CRAC kinase activity, under conditions sufficient to allow the candidate agent to interact with the SOC/CRAC polypeptide and modulate its kinase activity;

b) detecting a kinase activity associated with the SOC/CRAC polypeptide in the presence of the candidate agent; and

c) comparing the kinase activity of step (b) with a control kinase activity of a SOC/CRAC polypeptide in the absence of the candidate agent to determine whether the candidate agent modulates SOC/CRAC kinase activity.

37. The method of claim 36, wherein the SOC/CRAC polypeptide comprises amino acids 999-1180 of the sequence represented as SEQ ID NO:24, or a fragment thereof that retains the kinase activity.

1/1

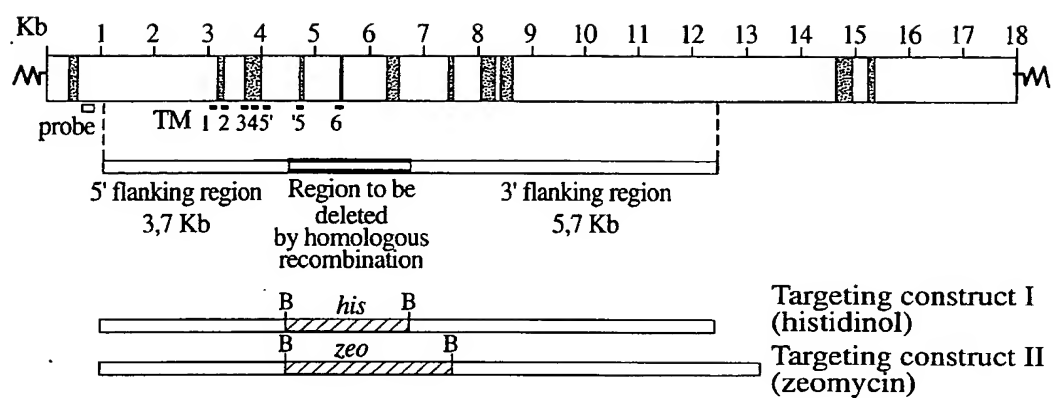


Fig. 1

-1-

SEQUENCE LISTING

<110> Beth Israel Deaconess Medical Center, Inc.
Scharenberg, Andrew

<120> CHARACTERIZATION OF A CALCIUM CHANNEL FAMILY

<130> B0662/7026WO/ERP/KA

<150> U.S. 60/114,220

<151> 1998-12-30

<150> U.S. 60/120,018

<151> 1999-01-29

<150> U.S. 60/140,415

<151> 1999-06-22

<160> 32

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 1212

<212> DNA

<213> Homo Sapiens

<400> 1

```
gcacgaggca aattttttgt tagtacacca tctcagccaa gttgcaaaaag ccacttgga      60
actggaacca aagatcaaga aactgtttgc tctaaagcta cagaaggaga taatacagaa      120
tttgagcat ttgtaggaca cagagatagc atggatttac agaggtttaa agaaacatca      180
aacaagataa aaatactatc caataacaat acttctgaaa acactttgaa acgagtgagt      240
tctcttgctg gatttactga ctgtcacaga acttccattc ctgttcattc aaaacaagaa      300
aaaatcagta gaagccatc taccgaagac actcatgaag tagattccaa agcagcttta      360
ataccggttt gtagatttca actaaacaga tatatattat taaatacatt aaactttttt      420
agataagatc tacaaagtgg tgatatttgg gactatatca aaaattcaaa aaaatttttc      480
ttaagaaaac tgacttttagc atagtagcag ttacagaaaa gtttcttaca gtgaatagtc      540
aggaatttta aagaaaaatt tatgcagaat aaaggcagga atctcttttt gtttgaattg      600
aagctaatta tatgaactca tttccagcta actgcgataa tgattgattt tgcaaatcc      660
ctttaaaagc acacactgac aagacaaaaa gctcaggaaa aggcagaaaa attactcctt      720
tataatcaag tattatatat aagtcagtgc tcataatttt gctcaagaaa atattgactt      780
acattcatat atatctgttc tggcatagag agattatggt gttaaaaatca tgttattgaa      840
aaaagtattt tcagtgggga aagaggttag ttaacaaaaga gattcacagt aacaaatcct      900
cctttctgga gggactcttc ctgaccctga gctgcacaac tttgcaacaa attaaagcct      960
aaccgaagat gacctcacia tggcaattta gaactcatgg gagtcaactt acataaacgg      1020
tatttgattt ctgataagat agtggaaatta ttggttatag atgacaaaat aagtatgttt      1080
aaagtgatga tggacataaa aaagttttta atataaaaca tgagaaaaaga aggagatact      1140
attcaaaaag actggcaaat ttgaaaaact agaaataaaa aaaaaaaaaa aaaatgagcg      1200
gccgcaagct tt                                     1212
```

<210> 2

<211> 141

<212> PRT

<213> Homo Sapiens

<400> 2

```
Ala Arg Gly Lys Phe Phe Val Ser Thr Pro Ser Gln Pro Ser Cys Lys
  1           5           10          15
Ser His Leu Glu Thr Gly Thr Lys Asp Gln Glu Thr Val Cys Ser Lys
          20          25          30
```

-2-

Ala	Thr	Glu	Gly	Asp	Asn	Thr	Glu	Phe	Gly	Ala	Phe	Val	Gly	His	Arg
	35						40					45			
Asp	Ser	Met	Asp	Leu	Gln	Arg	Phe	Lys	Glu	Thr	Ser	Asn	Lys	Ile	Lys
	50					55					60				
Ile	Leu	Ser	Asn	Asn	Asn	Thr	Ser	Glu	Asn	Thr	Leu	Lys	Arg	Val	Ser
	65				70					75				80	
Ser	Leu	Ala	Gly	Phe	Thr	Asp	Cys	His	Arg	Thr	Ser	Ile	Pro	Val	His
			85						90					95	
Ser	Lys	Gln	Glu	Lys	Ile	Ser	Arg	Arg	Pro	Ser	Thr	Glu	Asp	Thr	His
		100						105					110		
Glu	Val	Asp	Ser	Lys	Ala	Ala	Leu	Ile	Pro	Val	Cys	Arg	Phe	Gln	Leu
		115					120					125			
Asn	Arg	Tyr	Ile	Leu	Leu	Asn	Thr	Leu	Asn	Phe	Phe	Arg			
	130					135					140				

<210> 3
 <211> 739
 <212> DNA
 <213> Homo Sapiens

<220>
 <221> unsure
 <222> (5)...(5)
 <223> UNKNOWN

<221> unsure
 <222> (21)...(22)
 <223> UNKNOWN

<221> unsure
 <222> (29)...(29)
 <223> UNKNOWN

<400> 3	
tcgantaggg gtcttccacc nncatactng gatgatgggt ggtgaagtct atgcatacga	60
aattgatgtg tgtgcaaacg attctgttat ccctcaaatc tgtggtcctg ggacgtgggt	120
gactccattt cttcaagcag tctacctctt tgwacagtat atcattatgg ttaatcttct	180
tattgcattt ytcaacaatg tgtatttaca agtgaaggca atttccaata ttgyatggaa	240
gtaccagcgt tatcatthta ttatggctta tcatgagaaa ccagttctgc ctcctccact	300
tatcattctt agccatatag tttctctggt ttgctgcata tgtaagagaa gaaagaaaaga	360
taagacttcc gatggaccaa aacttttctt aacagaagaa gatcaaaaaga aacttcatga	420
ttttgaagag cagtgtgttg aaatgtattt caatgaaaaa gatgacaaat ttcattctgg	480
gagtgaagag agaattcgtg tcacttttga aagagtggaa cagatgtgca ttcagattaa	540
agaagttgga gatccgtgtc aactacataa aaagatcatt acaatcatta gattctcaaa	600
ttggccattt gcaagatctt tcagccctga cggtagatac attaaaaaca ctcactggcc	660
aaaagcgtcg gaagctagca aagttcataa tgaaatcaca cgagaactga gcatttccaa	720
acacttggct caaaacctt	739

<210> 4
 <211> 235
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> UNSURE
 <222> (41)...(41)
 <223> UNKNOWN

<221> UNSURE
 <222> (54)...(54)

-3-

<223> UNKNOWN

<221> UNSURE

<222> (68)...(68)

<223> UNKNOWN

<400> 4

Met	Met	Val	Gly	Glu	Val	Tyr	Ala	Tyr	Glu	Ile	Asp	Val	Cys	Ala	Asn
1				5					10					15	
Asp	Ser	Val	Ile	Pro	Gln	Ile	Cys	Gly	Pro	Gly	Thr	Trp	Leu	Thr	Pro
			20					25					30		
Phe	Leu	Gln	Ala	Val	Tyr	Leu	Phe	Xaa	Gln	Tyr	Ile	Ile	Met	Val	Asn
		35					40					45			
Leu	Leu	Ile	Ala	Phe	Xaa	Asn	Asn	Val	Tyr	Leu	Gln	Val	Lys	Ala	Ile
		50				55					60				
Ser	Asn	Ile	Xaa	Trp	Lys	Tyr	Gln	Arg	Tyr	His	Phe	Ile	Met	Ala	Tyr
65					70					75					80
His	Glu	Lys	Pro	Val	Leu	Pro	Pro	Pro	Leu	Ile	Ile	Leu	Ser	His	Ile
				85					90					95	
Val	Ser	Leu	Phe	Cys	Cys	Ile	Cys	Lys	Arg	Arg	Lys	Lys	Asp	Lys	Thr
			100					105					110		
Ser	Asp	Gly	Pro	Lys	Leu	Phe	Leu	Thr	Glu	Glu	Asp	Gln	Lys	Lys	Leu
		115					120					125			
His	Asp	Phe	Glu	Glu	Gln	Cys	Val	Glu	Met	Tyr	Phe	Asn	Glu	Lys	Asp
		130				135						140			
Asp	Lys	Phe	His	Ser	Gly	Ser	Glu	Glu	Arg	Ile	Arg	Val	Thr	Phe	Glu
145					150					155					160
Arg	Val	Glu	Gln	Met	Cys	Ile	Gln	Ile	Lys	Glu	Val	Gly	Asp	Pro	Cys
				165					170					175	
Gln	Leu	His	Lys	Lys	Ile	Ile	Thr	Ile	Ile	Arg	Phe	Ser	Asn	Trp	Pro
			180					185					190		
Phe	Ala	Arg	Ser	Phe	Ser	Pro	Asp	Gly	Arg	Tyr	Ile	Lys	Asn	Thr	His
		195					200					205			
Trp	Pro	Lys	Ala	Ser	Glu	Ala	Ser	Lys	Val	His	Asn	Glu	Ile	Thr	Arg
		210				215					220				
Glu	Leu	Ser	Ile	Ser	Lys	His	Leu	Ala	Gln	Asn					
225					230					235					

<210> 5

<211> 1579

<212> DNA

<213> Homo Sapiens

<220>

<221> unsure

<222> (368)...(368)

<223> g or c

<221> unsure

<222> (372)...(372)

<223> g or c

<221> unsure

<222> (374)...(374)

<223> g or a

<221> unsure

<222> (375)...(375)

<223> g or c

-4-

<221> unsure
 <222> (387)... (387)

<221> unsure
 <222> (482)... (482)

<400> 5
 acgtcgccctg caggtaccgg tccggaattc ccggggtcgac ccacgcgtcc ggcatgggtgt 60
 tgtaaatata ctttagctoct ctcttcctca aggtgatctt gaaagtaata atcctttttca 120
 ttgtaatatatt ttaatgaaag atgacaaaga tccccagtggt aatatatttg gtcaagactt 180
 acctgcagta cccagagaa aagaatttaa ttttccagag gctgggttcct cttctgggtgc 240
 cttattccca agtgctggtt ccctccaga actgcgacag agactacatg gggtagaact 300
 cttaaaaata ttttaataaaa atcaaaaatt aggcagttca tctactagca taccacatct 360
 gtcatccsca csarscaaat tttttgntag tacaccatct cagccaagtt gcaaaagcca 420
 cttggaaact ggaaccaaaag atcaagaaac tgtttgctct aaagctacag aaggagataa 480
 tncagaatatt ggagcatatt taggacacag agatagcatg gatttacaga ggtttaaaga 540
 aacatcaaac aagataaaaa tactatccaa taacaatact tctgaaaaca ctttgaaacg 600
 agtgagttct cttgctggat ttactgactg tcacagaact tccattcctg ttcatcctaa 660
 acaagaaaaa atcagtagaa ggccatctac cgaagacact catgaagtag attccaaagc 720
 agctttaata ccggttttga gatttcaact aaacagatat atattattaa atacattaaa 780
 ctttttttaga taagatctac aaagtgggtga tatttgggac tatatcaaaa attcaaaaaa 840
 atttttctta agaaaactga ctttagcata gttagcagtt cagaaaagtt tcttacagtg 900
 aatagtcagg aatttttaaag aaaaatttat gcagaataaa ggcaggaatc tctttttggt 960
 tgattgaag ctaattatat gaactcattt ccagctaact gcgataatga ttgattttgc 1020
 aaattccctt taaaagcaca cactgacaag acaaaaagct caggaaaagg cagaaaaatt 1080
 actcctttat aatcaagtat tatatataag tcagtgctca taattttgct caagaaaata 1140
 ttgactttaca ttcatatata tctgttctgg catagagaga ttatgttggt aaaaatcatgt 1200
 tattgaaaaa agttatttca gtggggaaag aggttagtta acaaagagat tcacagtaac 1260
 aaatccctct tcttgagggg actcttcctg accctgagct gcacaacttt gcaacaaatt 1320
 aaagcctaac cgaagatgac ctcaaatgg caatttagaa ctcatgggag tcaacttaca 1380
 taaacgggtat ttgatttctg ataagatagt ggaattattg gttatagatg acaaaataag 1440
 tatgtttaaa gtgatgatgg acataaaaaa gttttaaata taaaacatga gaaaagaagg 1500
 agatactatt caaaaagact ggcaaatattg aaaaactaga aataaaaaaa aaaaaaaaaa 1560
 atgagcggcc gcaagcttt 1579

<210> 6
 <211> 243
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> UNSURE
 <222> (103)... (105)
 <223> UNKNOWN

<221> UNSURE
 <222> (109)... (109)
 <223> UNKNOWN

<221> UNSURE
 <222> (141)... (141)
 <223> UNKNOWN

<400> 6
 Val Asn Thr Leu Ser Ser Ser Leu Pro Gln Gly Asp Leu Glu Ser Asn
 1 5 10 15
 Asn Pro Phe His Cys Asn Ile Leu Met Lys Asp Asp Lys Asp Pro Gln
 20 25 30
 Cys Asn Ile Phe Gly Gln Asp Leu Pro Ala Val Pro Gln Arg Lys Glu
 35 40 45

-5-

Phe Asn Phe Pro Glu Ala Gly Ser Ser Ser Gly Ala Leu Phe Pro Ser
 50 55 60
 Ala Val Ser Pro Pro Glu Leu Arg Gln Arg Leu His Gly Val Glu Leu
 65 70 75 80
 Leu Lys Ile Phe Asn Lys Asn Gln Lys Leu Gly Ser Ser Ser Thr Ser
 85 90 95
 Ile Pro His Leu Ser Ser Xaa Xaa Xaa Lys Phe Phe Xaa Ser Thr Pro
 100 105 110
 Ser Gln Pro Ser Cys Lys Ser His Leu Glu Thr Gly Thr Lys Asp Gln
 115 120 125
 Glu Thr Val Cys Ser Lys Ala Thr Glu Gly Asp Asn Xaa Glu Phe Gly
 130 135 140
 Ala Phe Val Gly His Arg Asp Ser Met Asp Leu Gln Arg Phe Lys Glu
 145 150 155 160
 Thr Ser Asn Lys Ile Lys Ile Leu Ser Asn Asn Asn Thr Ser Glu Asn
 165 170 175
 Thr Leu Lys Arg Val Ser Ser Leu Ala Gly Phe Thr Asp Cys His Arg
 180 185 190
 Thr Ser Ile Pro Val His Ser Lys Gln Glu Lys Ile Ser Arg Arg Pro
 195 200 205
 Ser Thr Glu Asp Thr His Glu Val Asp Ser Lys Ala Ala Leu Ile Pro
 210 215 220
 Val Cys Arg Phe Gln Leu Asn Arg Tyr Ile Leu Leu Asn Thr Leu Asn
 225 230 235 240
 Phe Phe Arg

<210> 7
 <211> 3532
 <212> DNA
 <213> Mus Musculus

<220>
 <221> unsure
 <222> (2420)... (2420)
 <223> unknown

<221> unsure
 <222> (2434)... (2434)
 <223> unknown

<221> unsure
 <222> (2461)... (2461)
 <223> unknown

<221> unsure
 <222> (2466)... (2466)
 <223> unknown

<221> unsure
 <222> (2470)... (2470)
 <223> unknown

<400> 7
 attatggcct atcatgaaaa accagtcctg cctcctcctc ttatcatcct cagccatata 60
 gtttcactgt tttgctgtgt atgcaaaaaga agaaagaaaag ataagacttc cgatgggcca 120
 aaacttttct taacagaaga agatcaaaaag aaactccatg attttgaaga gcagtgtgtt 180
 gagatgtact ttgatgagaa agatgacaaa ttcaattctg ggagtgaaga gagaatccgg 240
 gtcacttttg aaagagtggg gcagatgagc attcagatta aagaagttgg agatcgtgtc 300
 aactacataa aaagatcatt acagtcttta gattctcaaa ttgggtcatct gcaagatctc 360

-6-

tcagccctaa	cagtagatac	attgaaaaca	cttacagccc	agaaagcttc	agaagctagt	420
aaagtgcaca	atgagatcac	acgagaattcg	agtattttcca	aacacttggc	tcagaatcctt	480
attgatgatg	ttcctgtaag	acctttgtgg	gaagaacctta	gtgctgtaaa	cacactgagt	540
tcctctcttc	ctcaagggtga	tcgggaaagt	aataatccctt	ttctttgtaa	tatttttatg	600
aaagatgaaa	aagaccccca	atataatctg	tttggacaag	atttggccgt	gatacccccag	660
agaaaagaat	tcaacattcc	agaggctggg	tcctcctgtg	gtgccttatt	cccaagtgtc	720
gtttctcccc	cagaattacg	acagagacga	catggggtag	aaatgttaaa	aatattttaat	780
aaaaatcaaa	aattaggcag	ttcacctaata	agttcaccac	atatgtcctc	cccaccaacc	840
aaatttttctg	tgagtacccc	atcccagcca	agttgcaaaa	gtcacttgga	atccacaacc	900
aaagatcaag	aaccctatttt	ctataaagct	gcagaagggg	ataacataga	atttggagca	960
tttgtgggac	acagagatag	tatggactta	cagagggtta	aagaaacatc	aaacaaaata	1020
agagaactgt	tatctaataga	tactcctgaa	aacactctga	aacatgtggg	tgctgtctgga	1080
tatagtgaat	ttgttaagac	ttctacttct	cttcactcgg	tgcaagcaga	aagctgtagt	1140
agaagagcgt	cgacggaaga	ctctccagaa	gtcgattcta	aagcagcttt	gttaccggat	1200
tggttacgag	atagaccatc	aaacagagaa	atgccatctg	aaggaggaac	attaaatggt	1260
cttgcttctc	catttaagcc	cgttttggat	acaaattact	attattcagc	tgtggaaga	1320
aataacctga	tgaggttgtc	acagagtatt	cccttcgttc	ctgtacctcc	acgagcgag	1380
cctgtcacag	tgtaccgtct	ggaggagagt	tctcccagta	tactgaataa	cagcatgtct	1440
tcattggtctc	agctaggcct	ctgtgccaaa	attgagtttt	taagtaaaga	ggaaatggaa	1500
ggtgggtttac	gaagagcagt	caaagtgtctg	tgtacctggg	cagagcacga	tatcctgaag	1560
tcagggcac	tctatatcat	taagtcattt	cttcctgagg	tgataaacac	atggtcaagc	1620
atttataaag	aagatacggg	tctacatctc	tgtctcagag	aaatacaaca	acagagcgca	1680
gcacaaaagc	tcacatttgc	ctttaatcag	atgaaaccca	aatccatacc	atatttctcca	1740
aggttccttg	aagttttcct	gttgactgac	cattcagcag	ggcagtggtt	tgctgtagaa	1800
gagtgcata	ctggtgaatt	tagaaaaata	aacaacaata	atggtgatga	aatcattcct	1860
acaaataactc	tagaagagat	catgctagcc	tttagccact	ggacctatga	atataccaga	1920
ggggagttac	tggtacttga	cttacaagga	gtgggagaaa	acttgactga	cccatctgta	1980
ataaaagctg	aagaaaaaag	atcctgtgac	atgggttttg	gccttgccaa	tctaggagaa	2040
gatgcaataa	aaaacttcaa	gagccaaaaca	tccactgtaa	ttcttgctgt	cgaaagctta	2100
aacttcccag	atttgaagag	gaatgactac	acgccttgga	taaaattata	tttctcagg	2160
atgagtcac	agatttgaat	cttcaatctg	gaaattccac	caaagaatca	gaagcaacaa	2220
attctgttcg	tctgatgtta	tagtgctgag	tcattgggtt	ttgcctacac	ttcacaaaag	2280
tgtaactgtc	agttttccct	tcgggggaat	tgatgatata	ggaagatgtg	tgcaaaatga	2340
gcttgctggc	cccacacata	gtctagaggt	aatgttctca	ttgaaaaacg	cctggagggtg	2400
gaggctgcag	atgccagtg	aaagtgtctg	ctgncagaga	gtcagtgctc	tcgggctggg	2460
naaggncggg	acccttgctg	ctgagagtgg	tggttctctt	cacctgggtc	aggaccattt	2520
accaaagtca	agtcttcaga	tttgattggc	tgctcagtc	cagcccattc	agctaaggaa	2580
actaaattgc	gcagcttttt	aaatggctga	agtcttctct	agtttgtgct	ctatgataat	2640
gatgttagct	ctcaactagg	tgtttggtgc	cacgggagaa	ctactcctta	caattttgct	2700
tcacaggcat	gttacaagac	ctgcactgaa	aaccgtttgt	cttccctctc	ttccctccctc	2760
ttttccctgt	agtattgagg	atcaaaccga	gggcctcatg	aagaccattt	tctaagagac	2820
attttattta	agaatcaact	atagagtcta	tgtttatgga	tacagccagt	ttttgttaaa	2880
caaaacctga	attgtgcaaa	agggtttttt	aacattttatc	aatgttaagt	aaaagaaagc	2940
catgataaat	aagaattaac	tcactgttca	atgggtgttt	cctgtgagga	aggttacagt	3000
tgtaacagcc	tgcaattgca	tacatctcca	aagattttaca	gacttagtgt	atcaaatcag	3060
agtgtcatgt	gagctctcac	attgaaaatt	ctataggaat	gtgtcaatgt	gaattctatt	3120
tctggtactt	aagaaatcag	ttgttggtat	atccttatac	agtataggga	gatcacaata	3180
caactttatg	ccaataaaat	ctaacttaat	tgcccagata	tttttgcata	tttagcaaca	3240
agaaaagctt	atcattttgac	tcaagtttta	tgctttctct	ttcttttcat	ttcctaggta	3300
ctaattttta	tttttatttg	gaaggagcag	tgtaaagctt	acttgtattc	aatagtgtat	3360
ctcatagata	cagacaaggc	cgcagagata	agctgttaaa	tagtgtttaa	tgttgatgtg	3420
gagagaaagg	tgtattactt	aaaaatacta	taccatatac	gttttgtata	tcattaaatc	3480
tttaaaagaa	attaaattta	ttcttggtta	aaaaaaaaaa	aaaaaaaaaa	aa	3532

<210> 8

<211> 475

<212> PRT

<213> Mus Musculus

<400> 8

-7-

Ile	Met	Ala	Tyr	His	Glu	Lys	Pro	Val	Leu	Pro	Pro	Pro	Leu	Ile	Ile
1				5					10					15	
Leu	Ser	His	Ile	Val	Ser	Leu	Phe	Cys	Cys	Val	Cys	Lys	Arg	Arg	Lys
			20					25					30		
Lys	Asp	Lys	Thr	Ser	Asp	Gly	Pro	Lys	Leu	Phe	Leu	Thr	Glu	Glu	Asp
		35				40						45			
Gln	Lys	Lys	Leu	His	Asp	Phe	Glu	Glu	Gln	Cys	Val	Glu	Met	Tyr	Phe
	50					55					60				
Asp	Glu	Lys	Asp	Asp	Lys	Phe	Asn	Ser	Gly	Ser	Glu	Glu	Arg	Ile	Arg
65					70				75					80	
Val	Thr	Phe	Glu	Arg	Val	Glu	Gln	Met	Ser	Ile	Gln	Ile	Lys	Glu	Val
				85					90					95	
Gly	Asp	Arg	Val	Asn	Tyr	Ile	Lys	Arg	Ser	Leu	Gln	Ser	Leu	Asp	Ser
			100					105					110		
Gln	Ile	Gly	His	Leu	Gln	Asp	Leu	Ser	Ala	Leu	Thr	Val	Asp	Thr	Leu
		115				120						125			
Lys	Thr	Lys	Thr	Ala	Gln	Lys	Ala	Ser	Glu	Ala	Ser	Lys	Val	His	Asn
	130					135					140				
Glu	Ile	Thr	Arg	Glu	Leu	Ser	Ile	Ser	Lys	His	Leu	Ala	Gln	Asn	Leu
145					150					155					160
Ile	Asp	Asp	Val	Pro	Val	Arg	Pro	Leu	Trp	Glu	Glu	Pro	Ser	Ala	Val
				165					170					175	
Asn	Thr	Leu	Ser	Ser	Ser	Leu	Pro	Gln	Gly	Asp	Arg	Glu	Ser	Asn	Asn
		180						185						190	
Pro	Phe	Leu	Cys	Asn	Ile	Phe	Met	Lys	Asp	Glu	Lys	Asp	Pro	Gln	Tyr
		195					200					205			
Asn	Leu	Phe	Gly	Gln	Asp	Leu	Pro	Val	Ile	Pro	Gln	Arg	Lys	Glu	Phe
	210					215					220				
Asn	Ile	Pro	Glu	Ala	Gly	Ser	Ser	Cys	Gly	Ala	Leu	Phe	Pro	Ser	Ala
225					230					235					240
Val	Ser	Pro	Pro	Glu	Leu	Arg	Gln	Arg	Arg	His	Gly	Val	Glu	Met	Leu
				245					250					255	
Lys	Ile	Phe	Asn	Lys	Asn	Gln	Lys	Leu	Gly	Ser	Ser	Pro	Asn	Ser	Ser
			260					265					270		
Pro	His	Met	Ser	Ser	Pro	Pro	Thr	Lys	Phe	Ser	Val	Ser	Thr	Pro	Ser
		275					280						285		
Gln	Pro	Ser	Cys	Lys	Ser	His	Leu	Glu	Ser	Thr	Thr	Lys	Asp	Gln	Glu
	290					295					300				
Pro	Ile	Phe	Tyr	Lys	Ala	Ala	Glu	Gly	Asp	Asn	Ile	Glu	Phe	Gly	Ala
305					310					315					320
Phe	Val	Gly	His	Arg	Asp	Ser	Met	Asp	Leu	Gln	Arg	Phe	Lys	Glu	Thr
				325					330					335	
Ser	Asn	Lys	Ile	Arg	Glu	Leu	Leu	Ser	Asn	Asp	Thr	Pro	Glu	Asn	Thr
			340					345					350		
Leu	Lys	His	Val	Gly	Ala	Ala	Gly	Tyr	Ser	Glu	Cys	Cys	Lys	Thr	Ser
		355					360					365			
Thr	Ser	Leu	His	Ser	Val	Gln	Ala	Glu	Ser	Cys	Ser	Arg	Arg	Ala	Ser
	370					375					380				
Thr	Glu	Asp	Ser	Pro	Glu	Val	Asp	Ser	Lys	Ala	Ala	Leu	Leu	Pro	Asp
385					390					395					400
Trp	Leu	Arg	Asp	Arg	Pro	Ser	Asn	Arg	Glu	Met	Pro	Ser	Glu	Gly	Gly
				405					410					415	
Thr	Leu	Asn	Gly	Leu	Ala	Ser	Pro	Phe	Lys	Pro	Val	Leu	Asp	Thr	Asn
			420					425					430		
Tyr	Tyr	Tyr	Ser	Ala	Val	Glu	Arg	Asn	Asn	Leu	Met	Arg	Leu	Ser	Gln
		435					440					445			
Ser	Ile	Pro	Phe	Val	Pro	Val	Pro	Pro	Arg	Gly	Glu	Pro	Val	Thr	Val
	450					455					460				
Tyr	Pro	Ser	Gly	Gly	Arg	Val	Leu	Pro	Val	Tyr					
465					470					475					

<210> 9
 <211> 5433
 <212> DNA
 <213> Mus Musculus

<220>
 <221> unsure
 <222> (5094)... (5094)
 <223> unknown

<400> 9

ggctgaaaga	gacctgagctg	tgcctctcca	ttccactgct	gtggcagggg	cagaaatctt	60
ggatagagaa	aaccttttgc	aaacgggaat	gtatctttgt	aattcctagc	acgaaagact	120
ctaacagggtg	ttgctgtggc	cagttcacca	accagcatat	ccccctctg	ccaagtgc	180
cacccagcaa	aaatgaagag	gaaagcaaac	aggtggagac	tcagcctgag	aaatggctctg	240
ttgccaagca	cacccagagc	taccaaacag	attcctatgg	agttcttgaa	ttccaggggtg	300
gcggatattc	caataaagcc	atgtatatcc	gtgtatccta	tgacaccaag	ccagactcac	360
tgctccatct	catggtgaaa	gattggcagc	tggaaactccc	caagctctta	atatctgtgc	420
atggaggcct	ccagaacttt	gagatgcagc	ccaagctgaa	acaagtcttt	gggaaaggcc	480
tgatcaaggc	tgctatgacc	accggggcct	ggatcttcac	cgggggtgtc	agcacagggtg	540
ttatcagcca	cgtaggggat	gccttgaaag	accactctc	caagtccaga	ggccgggttt	600
gtgctatagg	aattgctcca	tggggcatcg	tggagaataa	ggaagacctg	gttggaagg	660
atgtaacaag	agtgtaccag	accatgtcca	accctctaag	taagctctct	gtgctcaaca	720
actccacac	ccacttcac	ctggctgaca	atggcaccct	gggcaagtat	ggcgccgagg	780
tgaagctgcg	aaggctgctg	gaaaagcaca	tctccctoca	gaagatcaac	acaagactgg	840
ggcagggcgt	gcccctcggt	ggtctcggtg	tggagggggg	ccctaactgt	gtgtccatcg	900
tcttggaata	cctgcaagaa	gagcctccca	tccctgtggt	gatttgtgat	ggcagcggac	960
gtgcctcgga	catcctgtcc	tttgccgaca	agtagctgtg	agaaggcgga	ataataaatg	1020
agtcctctcg	ggagcagctt	ctagttacca	ttcagaaaac	atttaattat	aataaggcac	1080
aatcacatca	gctgttttga	attataatgg	agtgcataaa	gaagaaagaa	ctcgtcactg	1140
tggttcagaat	gggttctgag	ggccagcagg	acatcgagat	ggcaatttta	actgccctgc	1200
tgaaggaac	aaacgtatct	gctccagatc	agctgagctt	ggcactgggt	tggaaaccgg	1260
tggacatagc	acgaagccag	atctttgtct	tggggcccca	ctggacgccc	ctgggaagcc	1320
tggcaccccc	acgggacagc	aaagccacgg	agaaggagaa	gaagccaccc	atggccacca	1380
ccaagggagg	aagaggaaaa	gggaaaggca	agaagaaagg	gaaagtgaag	gaggaagtgg	1440
aggaagaaac	tgacccccgg	aagatagagc	tgctgaactg	ggtgaatgct	ttggaagcaag	1500
cgatgctaga	tgcttttagtc	ttagatcggt	tcgactttgt	gaagctcctg	attgaaaacg	1560
gagtgaacat	gcaacacttt	ctgaccattc	cgaaggctga	ggagctctat	aacacaagac	1620
tgggtccacc	aaacacactt	catctgctgg	tgagggatgt	gaaaaagagc	aaccttccgc	1680
ctgattacca	catcagcctc	atagacatcg	ggctcgtgct	ggagtacctc	atgggaggag	1740
cctaccgctg	caactacact	cggaaaaaact	ttcggaccct	ttacaacaac	ttgtttggac	1800
caaagaggcc	taaagctctt	aaacttctgg	gaatggaaga	tgatgagcct	ccagctaaag	1860
ggaagaaaaa	aaaaaaaaag	aaaaaggagg	aagagatcga	cattgatgtg	gacgaccctg	1920
ccgtgagtcg	gttccagtat	cccttccacg	agctgatggg	gtgggcagtg	ctgatgaaac	1980
gccagaaaat	ggcagtggtc	ctctggcagc	gaggggaaga	gagcatggcc	aaggccctgg	2040
tggcctgcaa	gctctacaag	gccatggccc	acgagtcctc	cgagagtgat	ctggtgggatg	2100
acatctccca	ggacttggat	aacaattcca	aagacttcgg	ccagcttgct	ttggagttat	2160
tagaccagtc	ctataagcat	gacgagcaga	tcgctatgaa	actcctgacc	tacgagctga	2220
aaaactggag	caactcgacc	tgctcacaac	tggcgtgggc	agccaaacac	cgggacttca	2280
ttgctcacac	ctgcagccag	atgctgctga	ccgatatgtg	gatgggaaga	ctgcggatgc	2340
ggaagaacct	cggcctgaag	gttatcatgg	ggattcttct	acccccacc	atcttgtttt	2400
tggaaatttcg	cacatatgat	gatttctcgt	atcaaacatc	caaggaaaac	gaggatggca	2460
aagaaaaaga	agaggaaaat	acggatgcaa	atgcagatgc	tggctcaaga	aagggggatg	2520
aggagaacga	gcataaaaaa	cagagaagta	ttcccatcgg	aacaaagatc	tgtgaattct	2580
ataacgcgcc	cattgtcaag	ttctgggttt	acacaatatc	atacttgggc	tacctgtctgc	2640
tgtttaacta	cgctatcctg	gtgcggatgg	atggctggcc	gtccctccag	gagtggatcg	2700
tcattctcta	catcgtgagc	ctggcggttag	agaagatacg	agagatcctc	atgtcagaac	2760
caggcaaaact	cagccagaaa	atcaaagttt	ggcttcagga	gtactggaac	atcacagatc	2820
tcgtggccat	ttccacattc	atgattggag	caattcttcg	cctacagaac	cagccctaca	2880

-9-

```

tgggctatgg ccgggtgata tactgtgtgg atatcatctt ctggtacatc cgtgtcctgg 2940
acatcttttg tgtaacaag tatctggggc catacgtgat gatgattgga aagatgatga 3000
tcgacatgct gtactttgtg gtcacatgct tggctgtgct catgagtttc ggagtagccc 3060
gtcaagccat tctgcatcca gaggagaagc cctcttgga actggcccga aacatcttct 3120
acatgcccta ctggatgata tatggagagg tgtttgcaga ccagatagac ctctacgcca 3180
tgaaattaa tcctccttgt ggtgagaacc tatatgatga ggaggccaag cggcttcctc 3240
cctgtatccc cggcgccctg ctcactccag cactcatggc gtgctatcta ctggtcgcca 3300
acatcctgct ggtgaacctg ctgattgctg tgttcaacaa tactttcttt gaagtaaaat 3360
caatatccaa ccaggtgtgg aagtccagc gatatcagct gattatgaca ttcatgaca 3420
ggccagtcct gccccaccg atgatcattt taagccacat ctacatcatc attatgcgtc 3480
tcagcggccg ctgcaggaaa aagagagaag ggaaccaaga ggaacgggat cgtggattga 3540
agctcttccct tagcgacgag gacgtaaaag ggctgcatga gttcaggag cagtgcgtgc 3600
aggagcactt ccgggagaag gaggatgagc agcagtcgtc cagcgacgag cgcacccggg 3660
tcacttctga aagagttgaa aatatgtcaa tgaggttggg agaaatcaat gaaagagaaa 3720
cttttatgaa aacttccctg cagactgttg accttcgact tgctcagcta gaagaattat 3780
ctaacagaat ggtgaatgct cttgaaaatc ttgcgggaat cgacaggtct gacctgatcc 3840
aggcacggtc ccgggcttct tctgaatgtg aggcaacgta tcttctccgg caaagcagca 3900
tcaatagcgc tgatggctac agcttgatc gatatcattt taacggagaa gagttattat 3960
ttgaggatac atctctctcc acgtcaccag ggacaggagt caggaaaaaa acctgttcc 4020
tccgtataaa ggaagagaag gacgtgaaa cgcacctagt ccagaaatgt cagaacagtc 4080
ttcacctttc actgggcaca agcacatcag caaccccaga tggcagtcac ctgacagtag 4140
atgacttaaa gaacgctgaa gactcaaaat taggtccaga tattgggatt tcaaagggaag 4200
atgatgaaag acagacagac tctaaaaaag aagaaactat ttccccaagt ttaataaaaa 4260
cagatgtgat acatggacag gacaaatcag atgttcaaaa cactcagcta acagtggaaa 4320
cgacaaatat agaaggcact atttccatc ccctggaaga aaccaaattt acacgtatt 4380
tccccgatga aacgatcaat gcttgtaaaa caatgaagtc cagaagcttc gtctattccc 4440
ggggaagaaa gctggctcgt ggggttaacc aggatgtaga gtacagttca atcacggacc 4500
agcaattgac gacggaatgg caatgccaag ttcaaaagat cacgcgtctc catagcacag 4560
atattcctta atctgtgtc gaagctgcag tgcaagctga gcaaaaagag cagtttgcag 4620
atatgcaaga tgaacacat gtcgctgaag caattcctcg aatccctcgc ttgtccctaa 4680
ccattactga cagaaatggg atggaaaact tactgtctgt gaagccagat caaactttgg 4740
gattcccatc tctcaggtca aaaagtttac atggacatcc taggaatgtg aaatccattc 4800
agggaaagtt agacagatct ggacatgccg tagtgtaag cagcttagta attgtgtctg 4860
gaatgacagc agaagaaaaa aagggttaaga aagagaaaagc ttccacagaa actgaatgct 4920
agtctgtttt gtttctttta ttttttttt taacagtcag aaacccta atgggtgtca 4980
tcttgcccca tctaaacac atmtccaatt tcttaaaaac attttccctt aaaaaatttt 5040
ggaaattcag acttgattta caatttaatg cactaaaagt agtattttgt tagnatatgt 5100
tagtaggctt agttttttca gttgcagtag tatcaaatga aagtgatgat actgtaacga 5160
agataaattg gctaatacgt atacaagatt atacaatctc tttattactg agggccacca 5220
aatagcctag gaagtgcct cgagcactga agtcaccatt agtcactca agaagtaagc 5280
aactagctgg gcacagtggc tcatgcctgt aatcctagca ctttgggagg ccaaggcaga 5340
aagatagctt gagtccagga gtttgagacc agcctgggca acatagtgat accccatctc 5400
ttaaaaaaaa aaaaaaaa ctgccctcgt gcc 5433

```

<210> 10

<211> 1533

<212> PRT

<213> Mus Musculus

<400> 10

```

Met Tyr Ile Arg Val Ser Tyr Asp Thr Lys Pro Asp Ser Leu Leu His
1          5          10          15
Leu Met Val Lys Asp Trp Gln Leu Glu Leu Pro Lys Leu Leu Ile Ser
20          25          30
Val His Gly Gly Leu Gln Asn Phe Glu Met Gln Pro Lys Leu Lys Gln
35          40          45
Val Phe Gly Lys Gly Leu Ile Lys Ala Ala Met Thr Thr Gly Ala Trp
50          55          60

```

-10-

Ile	Phe	Thr	Gly	Gly	Val	Ser	Thr	Gly	Val	Ile	Ser	His	Val	Gly	Asp
65					70					75					80
Ala	Leu	Lys	Asp	His	Ser	Ser	Lys	Ser	Arg	Gly	Arg	Val	Cys	Ala	Ile
				85					90					95	
Gly	Ile	Ala	Pro	Trp	Gly	Ile	Val	Glu	Asn	Lys	Glu	Asp	Leu	Val	Gly
			100					105					110		
Lys	Asp	Val	Thr	Arg	Val	Tyr	Gln	Thr	Met	Ser	Asn	Pro	Leu	Ser	Lys
			115				120					125			
Leu	Ser	Val	Leu	Asn	Asn	Ser	His	Thr	His	Phe	Ile	Leu	Ala	Asp	Asn
			130			135					140				
Gly	Thr	Leu	Gly	Lys	Tyr	Gly	Ala	Glu	Val	Lys	Leu	Arg	Arg	Leu	Leu
145					150					155					160
Glu	Lys	His	Ile	Ser	Leu	Gln	Lys	Ile	Asn	Thr	Arg	Leu	Gly	Gln	Gly
				165					170					175	
Val	Pro	Leu	Val	Gly	Leu	Val	Val	Glu	Gly	Gly	Pro	Asn	Val	Val	Ser
			180					185					190		
Ile	Val	Leu	Glu	Tyr	Leu	Gln	Glu	Glu	Pro	Pro	Ile	Pro	Val	Val	Ile
			195			200						205			
Cys	Asp	Gly	Ser	Gly	Arg	Ala	Ser	Asp	Ile	Leu	Ser	Phe	Ala	His	Lys
	210					215					220				
Tyr	Cys	Glu	Glu	Gly	Gly	Ile	Ile	Asn	Glu	Ser	Leu	Arg	Glu	Gln	Leu
225					230					235					240
Leu	Val	Thr	Ile	Gln	Lys	Thr	Phe	Asn	Tyr	Asn	Lys	Ala	Gln	Ser	His
				245				250						255	
Gln	Leu	Phe	Ala	Ile	Ile	Met	Glu	Cys	Met	Lys	Lys	Lys	Glu	Leu	Val
			260				265						270		
Thr	Val	Phe	Arg	Met	Gly	Ser	Glu	Gly	Gln	Gln	Asp	Ile	Glu	Met	Ala
			275				280					285			
Ile	Leu	Thr	Ala	Leu	Leu	Lys	Gly	Thr	Asn	Val	Ser	Ala	Pro	Asp	Gln
					290		295				300				
Leu	Ser	Leu	Ala	Leu	Ala	Trp	Asn	Arg	Val	Asp	Ile	Ala	Arg	Ser	Gln
305					310					315					320
Ile	Phe	Val	Phe	Gly	Pro	His	Trp	Thr	Pro	Leu	Gly	Ser	Leu	Ala	Pro
				325					330					335	
Pro	Thr	Asp	Ser	Lys	Ala	Thr	Glu	Lys	Glu	Lys	Lys	Pro	Pro	Met	Ala
			340					345					350		
Thr	Thr	Lys	Gly	Gly	Arg	Gly	Lys	Gly	Lys	Gly	Lys	Lys	Lys	Gly	Lys
		355					360					365			
Val	Lys	Glu	Glu	Val	Glu	Glu	Glu	Thr	Asp	Pro	Arg	Lys	Ile	Glu	Leu
					370		375				380				
Leu	Asn	Trp	Val	Asn	Ala	Leu	Glu	Gln	Ala	Met	Leu	Asp	Ala	Leu	Val
385					390					395					400
Leu	Asp	Arg	Val	Asp	Phe	Val	Lys	Leu	Leu	Ile	Glu	Asn	Gly	Val	Asn
				405				410						415	
Met	Gln	His	Phe	Leu	Thr	Ile	Pro	Arg	Leu	Glu	Glu	Leu	Tyr	Asn	Thr
			420				425						430		
Arg	Leu	Gly	Pro	Pro	Asn	Thr	Leu	His	Leu	Leu	Val	Arg	Asp	Val	Lys
			435				440					445			
Lys	Ser	Asn	Leu	Pro	Pro	Asp	Tyr	His	Ile	Ser	Leu	Ile	Asp	Ile	Gly
			450			455					460				
Leu	Val	Leu	Glu	Tyr	Leu	Met	Gly	Gly	Ala	Tyr	Arg	Cys	Asn	Tyr	Thr
465					470					475					480
Arg	Lys	Asn	Phe	Arg	Thr	Leu	Tyr	Asn	Asn	Leu	Phe	Gly	Pro	Lys	Arg
				485				490						495	
Pro	Lys	Ala	Leu	Lys	Leu	Leu	Gly	Met	Glu	Asp	Asp	Glu	Pro	Pro	Ala
			500					505					510		
Lys	Gly	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Lys	Glu	Glu	Glu	Ile	Asp	Ile
			515				520					525			
Asp	Val	Asp	Asp	Pro	Ala	Val	Ser	Arg	Phe	Gln	Tyr	Pro	Phe	His	Glu
			530			535					540				

-11-

Leu	Met	Val	Trp	Ala	Val	Leu	Met	Lys	Arg	Gln	Lys	Met	Ala	Val	Phe
545					550					555					560
Leu	Trp	Gln	Arg	Gly	Glu	Glu	Ser	Met	Ala	Lys	Ala	Leu	Val	Ala	Cys
				565					570					575	
Lys	Leu	Tyr	Lys	Ala	Met	Ala	His	Glu	Ser	Ser	Glu	Ser	Asp	Leu	Val
			580					585					590		
Asp	Asp	Ile	Ser	Gln	Asp	Leu	Asp	Asn	Asn	Ser	Lys	Asp	Phe	Gly	Gln
		595					600					605			
Leu	Ala	Leu	Glu	Leu	Leu	Asp	Gln	Ser	Tyr	Lys	His	Asp	Glu	Gln	Ile
	610					615					620				
Ala	Met	Lys	Leu	Leu	Thr	Tyr	Glu	Leu	Lys	Asn	Trp	Ser	Asn	Ser	Thr
625					630					635					640
Cys	Leu	Lys	Leu	Ala	Val	Ala	Ala	Lys	His	Arg	Asp	Phe	Ile	Ala	His
				645					650					655	
Thr	Cys	Ser	Gln	Met	Leu	Leu	Thr	Asp	Met	Trp	Met	Gly	Arg	Leu	Arg
			660					665					670		
Met	Arg	Lys	Asn	Pro	Gly	Leu	Lys	Val	Ile	Met	Gly	Ile	Leu	Leu	Pro
		675					680					685			
Pro	Thr	Ile	Leu	Phe	Leu	Glu	Phe	Arg	Thr	Tyr	Asp	Asp	Phe	Ser	Tyr
	690					695					700				
Gln	Thr	Ser	Lys	Glu	Asn	Glu	Asp	Gly	Lys	Glu	Lys	Glu	Glu	Glu	Asn
705					710					715					720
Thr	Asp	Ala	Asn	Ala	Asp	Ala	Gly	Ser	Arg	Lys	Gly	Asp	Glu	Glu	Asn
				725					730					735	
Glu	His	Lys	Lys	Gln	Arg	Ser	Ile	Pro	Ile	Gly	Thr	Lys	Ile	Cys	Glu
			740					745					750		
Phe	Tyr	Asn	Ala	Pro	Ile	Val	Lys	Phe	Trp	Phe	Tyr	Thr	Ile	Ser	Tyr
		755					760					765			
Leu	Gly	Tyr	Leu	Leu	Leu	Phe	Asn	Tyr	Val	Ile	Leu	Val	Arg	Met	Asp
	770					775					780				
Gly	Trp	Pro	Ser	Leu	Gln	Glu	Trp	Ile	Val	Ile	Ser	Tyr	Ile	Val	Ser
785					790					795					800
Leu	Ala	Leu	Glu	Lys	Ile	Arg	Glu	Ile	Leu	Met	Ser	Glu	Pro	Gly	Lys
				805					810					815	
Leu	Ser	Gln	Lys	Ile	Lys	Val	Trp	Leu	Gln	Glu	Tyr	Trp	Asn	Ile	Thr
			820					825					830		
Asp	Leu	Val	Ala	Ile	Ser	Thr	Phe	Met	Ile	Gly	Ala	Ile	Leu	Arg	Leu
		835					840					845			
Gln	Asn	Gln	Pro	Tyr	Met	Gly	Tyr	Gly	Arg	Val	Ile	Tyr	Cys	Val	Asp
	850					855					860				
Ile	Ile	Phe	Trp	Tyr	Ile	Arg	Val	Leu	Asp	Ile	Phe	Gly	Val	Asn	Lys
865					870					875					880
Tyr	Leu	Gly	Pro	Tyr	Val	Met	Met	Ile	Gly	Lys	Met	Met	Ile	Asp	Met
				885					890					895	
Leu	Tyr	Phe	Val	Val	Ile	Met	Leu	Val	Val	Leu	Met	Ser	Phe	Gly	Val
			900					905					910		
Ala	Arg	Gln	Ala	Ile	Leu	His	Pro	Glu	Glu	Lys	Pro	Ser	Trp	Lys	Leu
			915				920					925			
Ala	Arg	Asn	Ile	Phe	Tyr	Met	Pro	Tyr	Trp	Met	Ile	Tyr	Gly	Glu	Val
						935					940				
Phe	Ala	Asp	Gln	Ile	Asp	Leu	Tyr	Ala	Met	Glu	Ile	Asn	Pro	Pro	Cys
945					950					955					960
Gly	Glu	Asn	Leu	Tyr	Asp	Glu	Glu	Gly	Lys	Arg	Leu	Pro	Pro	Cys	Ile
				965					970					975	
Pro	Gly	Ala	Trp	Leu	Thr	Pro	Ala	Leu	Met	Ala	Cys	Tyr	Leu	Leu	Val
			980					985					990		
Ala	Asn	Ile	Leu	Leu	Val	Asn	Leu	Leu	Ile	Ala	Val	Phe	Asn	Asn	Thr
		995					1000					1005			
Phe	Phe	Glu	Val	Lys	Ser	Ile	Ser	Asn	Gln	Val	Trp	Lys	Phe	Gln	Arg
	1010					1015					1020				

-12-

Tyr Gln Leu Ile Met Thr Phe His Asp Arg Pro Val Leu Pro Pro Pro
 1025 1030 1035 104
 Met Ile Ile Leu Ser His Ile Tyr Ile Ile Met Arg Leu Ser Gly
 1045 1050 1055
 Arg Cys Arg Lys Lys Arg Glu Gly Asp Gln Glu Glu Arg Asp Arg Gly
 1060 1065 1070
 Leu Lys Leu Phe Leu Ser Asp Glu Glu Leu Lys Arg Leu His Glu Phe
 1075 1080 1085
 Glu Glu Gln Cys Val Gln Glu His Phe Arg Glu Lys Glu Asp Glu Gln
 1090 1095 1100
 Gln Ser Ser Ser Asp Glu Arg Ile Arg Val Thr Ser Glu Arg Val Glu
 1105 1110 1115 112
 Asn Met Ser Met Arg Leu Glu Glu Ile Asn Glu Arg Glu Thr Phe Met
 1125 1130 1135
 Lys Thr Ser Leu Gln Thr Val Asp Leu Arg Leu Ala Gln Leu Glu Glu
 1140 1145 1150
 Leu Ser Asn Arg Met Val Asn Ala Leu Glu Asn Leu Ala Gly Ile Asp
 1155 1160 1165
 Arg Ser Asp Leu Ile Gln Ala Arg Ser Arg Ala Ser Ser Glu Cys Glu
 1170 1175 1180
 Ala Thr Tyr Leu Leu Arg Gln Ser Ser Ile Asn Ser Ala Asp Gly Tyr
 1185 1190 1195 120
 Ser Leu Tyr Arg Tyr His Phe Asn Gly Glu Glu Leu Leu Phe Glu Asp
 1205 1210 1215
 Thr Ser Leu Ser Thr Ser Pro Gly Thr Gly Val Arg Lys Lys Thr Cys
 1220 1225 1230
 Ser Phe Arg Ile Lys Glu Glu Lys Asp Val Lys Thr His Leu Val Pro
 1235 1240 1245
 Glu Cys Gln Asn Ser Leu His Leu Ser Leu Gly Thr Ser Thr Ser Ala
 1250 1255 1260
 Thr Pro Asp Gly Ser His Leu Ala Val Asp Asp Leu Lys Asn Ala Glu
 1265 1270 1275 128
 Glu Ser Lys Leu Gly Pro Asp Ile Gly Ile Ser Lys Glu Asp Asp Glu
 1285 1290 1295
 Arg Gln Thr Asp Ser Lys Lys Glu Glu Thr Ile Ser Pro Ser Leu Asn
 1300 1305 1310
 Lys Thr Asp Val Ile His Gly Gln Asp Lys Ser Asp Val Gln Asn Thr
 1315 1320 1325
 Gln Leu Thr Val Glu Thr Thr Asn Ile Glu Gly Thr Ile Ser Tyr Pro
 1330 1335 1340
 Leu Glu Glu Thr Lys Ile Thr Arg Tyr Phe Pro Asp Glu Thr Ile Asn
 1345 1350 1355 136
 Ala Cys Lys Thr Met Lys Ser Arg Ser Phe Val Tyr Ser Arg Gly Arg
 1365 1370 1375
 Lys Leu Val Gly Gly Val Asn Gln Asp Val Glu Tyr Ser Ser Ile Thr
 1380 1385 1390
 Asp Gln Gln Leu Thr Thr Glu Trp Gln Cys Gln Val Gln Lys Ile Thr
 1395 1400 1405
 Arg Ser His Ser Thr Asp Ile Pro Tyr Ile Val Ser Glu Ala Ala Val
 1410 1415 1420
 Gln Ala Glu Gln Lys Glu Gln Phe Ala Asp Met Gln Asp Glu His His
 1425 1430 1435 144
 Val Ala Glu Ala Ile Pro Arg Ile Pro Arg Leu Ser Leu Thr Ile Thr
 1445 1450 1455
 Asp Arg Asn Gly Met Glu Asn Leu Leu Ser Val Lys Pro Asp Gln Thr
 1460 1465 1470
 Leu Gly Phe Pro Ser Leu Arg Ser Lys Ser Leu His Gly His Pro Arg
 1475 1480 1485
 Asn Val Lys Ser Ile Gln Gly Lys Leu Asp Arg Ser Gly His Ala Ser
 1490 1495 1500

-13-

Ser Val Ser Ser Leu Val Ile Val Ser Gly Met Thr Ala Glu Glu Lys
 1505 1510 1515 152
 Lys Val Lys Lys Glu Lys Ala Ser Thr Glu Thr Glu Cys
 1525 1530

<210> 11
 <211> 6220
 <212> DNA
 <213> Homo Sapiens

<400> 11
 tgtgcagaat tgtacagttg cgaaccatg tcgctggcag ctggtgctgg cggaggagac 60
 ttccctgtgc ggtgctcagt gcatctgcac ccgtggggga gggagctctt tctctggccc 120
 tgcagtcacc tgaggttgtt accattatga acggccgctg ggacccccgc atgtgcatgt 180
 actccccag agtgtccggg ggccccagcc aaggacaca tctcacgcag ctgggaacat 240
 gtgcaggctg atgaagagaa ccggatgagg gcttcacatg aggaagcatg tggccaggtc 300
 ctctcagaac atcagcctca tcttcctgtc tctgatctat ttcaccaacc accccatgtg 360
 tctctagaac cccagtgtag cgagctggag agaggactgt cctgagggca gcaggcctgg 420
 ttgcagctgg cgtgggggtc tcagaatgga gccctcagcc ctgaggaaag ctggctcgga 480
 gcaggaggag ggctttgagg ggctgcccag aagggtcact gacctgggga tggctctcaa 540
 tctccggcgc agcaacagca gctcttcaa gagctggagg ctacagtgcc ccttcggcaa 600
 caatgacaag caagaaagcc tcagttcgtg gattcctgaa aacatcaaga agaaagaatg 660
 cgtgtatttt gtggaaagt ccaactgtc tgatgctggg aaggtggtgt gtcagtgtgg 720
 ctacacgcag gagcagcact tggaggaggc taccaagccc cacaccttc agggcacaca 780
 gtgggaccca aagaaacatg tccaggagat gccaacccat gcctttggcg acatcgtctt 840
 caccggcctg agccagaagg tgaaaaagta cgtccgagtc tcccaggaca cgccctccag 900
 cgtgatctac cacctcatga ccagcactg ggggctggac gtccccaatc tcttgatctc 960
 ggtgaccggg ggggccaaga acttcaacat gaagccgcgg ctgaagagca ttttccgcag 1020
 aggcctggtc aaggtggctc agaccacagg ggccctggatc atcacagggg ggtccacac 1080
 cggcgctcatg aagcaggtag gcgaggcgtt gcgggacttc agcctgagca gcagctacaa 1140
 ggaaggcgag ctcatcacca tcggagtcgc cacctggggc actgtccacc gccgcgaggg 1200
 cctgatccat cccacgggca gcttccccgc cgagtacata ctggatgagg atggccaagg 1260
 gaacctgacc tgcctagaca gcaaccactc tcaattcatc ctggtggacg acgggaccca 1320
 cggccagtac ggggtggaga ttctctctgag gaccaggctg gagaagttca tatcggagca 1380
 gaccaaggaa agaggaggtg tggccatcaa gatccccatc gtgtgcgtgg tgcaggagg 1440
 cggccccggc acgttgcaca ccatcgacaa cgccaccacc aacggcacc cctgtgtggt 1500
 tgtggagggc tcgggcccgc tggccgacgt cattgcccag gtggccaacc tgcctgtctc 1560
 ggacatcact atctccctga tccagcagaa actgagcgtg ttcttccagg agatgtttga 1620
 gaccttcacg gaaagcagga ttgtcagtg gacaaaaag atccaagata ttgtccggag 1680
 gcggcagctg ctgactgtct tccgggaagg caaggatggt cagcaggacg tggatgtggc 1740
 catcttgacg gccttgctga aagcctcacg gagccaagac cactttggcc acgagaactg 1800
 ggaccaccag ctgaaactgg cagtggcatg gaatcgcggt gacattgccc gcagtggat 1860
 cttcatggat gagtggcagt ggaagccttc agatctgcac cccacgatga cagctgcact 1920
 catctccaac aagcctgagt ttgtgaagct cttcctggaa aacggggtgc agctgaagga 1980
 gtttgtcacc tgggacacct tgctctacct gtacgagaac ctggacccct cctgcctgtt 2040
 ccacagcaag ctgcaaaaagg tgctggtgga ggatcccag cgcccggtt gcgcgcccgc 2100
 ggcgccccgc ctgcagatgc accacgtggc ccagggtgct cgggagctgc tgggggactt 2160
 caccgagccg ctttatcccc ggccccgca caacgaccgg ctgcggctcc tgcctgccgt 2220
 tccccacgtc aagctcaacg tgcaggaggt gagcctccgg tccctctaca agcgttcctc 2280
 aggccatgtg accttcacca tggaccccat ccgtgacctt ctcatgtgg ccatgtcca 2340
 gaaccgtcgg gagctggcag gaatcatctg ggctcagagc caggactgca tcgcagcggc 2400
 cttggcctgc agcaagatcc tgaaggaaact tgagcacaga gccatcgggg tcttcaccga 2460
 ggaggagatg ctggcgctgg cggaggagta gaaactgtc acccgctgt cagaggcctg 2520
 gtgctaccgg aaggacgaag agagagccca gagggccaa gacatgaagt ttgtgtctca 2580
 ggggaagacc acctgcctgc agctcgccct ggaggccaa cagctctccg tggacaatgg 2640
 cgggggcatc caggccttcc tgaccaaggt gtggtggggc ctctccgctg gcctcatctc 2700
 cttgtggcgt gtgaccctgt gcatgctggc cttcccgctg cctctcaccg gtgccttctt 2760
 cttcagggag aagaggctgc aggatgtggg caccgccgcg gcccgcgccc gtgccttctt 2820
 caccgcaccc gtggtgtgtc tccacctgaa catcctctcc tacttcgcct tccctgcct 2880
 gttcgcctac gtgctcatgg tggacttcca gcctgtgccc tccgtggtgc agtgtgccat 2940

-14-

ctacctctgg	ctcttctcct	tgggtgtgca	ggagatgcgg	cagctcttct	atgaccctga	3000
cgagtgcggg	ctgatgaaga	aggcagcctt	gtacttccagt	gacttctgga	ataagctgga	3060
cgtcggcgca	atcttgctct	tcgtggcagg	gctgacctgc	aggctcatcc	cggcgacgct	3120
gtaccccggg	cgcgtcatcc	tctctctgga	cttcatcctg	ttctgcctcc	ggctcatgca	3180
catttttacc	atcagtaaga	cgctggggcc	caagatcatc	attgtgaagc	ggatgatgaa	3240
ggacgtcttc	ttcttcctct	tcctgctggc	tgtgtgggtg	gtgtccttcg	gggtggccaa	3300
gcaggccatc	ctcatccaca	acgagcgccg	ggtggactgg	ctgttccgag	gggccgtcta	3360
ccactcctac	ctcaccatct	tcgggcagat	cccgggctac	atcgacggtg	tgaacttcaa	3420
cgcgtgagac	tgccagcccca	atggcaccca	cccctacaag	cctaagtgcc	ccgagagcga	3480
cgcgagcgag	cagaggccgg	ccttccctga	gtggctgacg	gtcctcctac	tctgcctcta	3540
cctgctcttc	accaacatcc	tgctgctcaa	cctcctcatc	gccatgttca	actacacctt	3600
ccagcagggtg	caggagcaca	cggaccagat	ttggaagtgc	cagcgccatg	acctgatcga	3660
ggagtaccac	ggcgcccccg	ccgcgcgcgc	ccccttcata	ctcctcagcc	acctgcagct	3720
cttcatcaag	agggtgggtc	tgaagactcc	ggccaagagg	cacaagcagc	tcaagaacaa	3780
gctggagaag	aacgaggagg	cggccctgct	atcctgggag	atctacctga	aggagaacta	3840
cctccagaac	cgacagttcc	agcaaaagca	gcggccccgag	cagaagatcg	aggacatcag	3900
caataagggt	gacgcatggg	tggacctgct	ggacctggac	ccactgaaga	ggtcgggctc	3960
catggagcag	agggtggcct	ccctggagga	gcagggtggcc	cagacagccc	gagccctgca	4020
ctggatcggtg	aggacgtgct	ggggccagcgg	cttcagctcg	gaggcggacg	tccccactct	4080
ggcctcccag	aaggccgcgg	aggagccgga	tgctgagccg	ggaggcagga	agaagacgga	4140
ggagccgggg	gacagctacc	acgtgaatgc	ccggcacctc	ctctacccca	actgccctgt	4200
cacgcgccttc	ccctggccca	acgagaaggt	gccctgggag	acggagttcc	tgatctatga	4260
cccacccttt	tacacggcag	agaggaagga	cgcgcccgcc	atggacccca	tgggagacac	4320
cctggagcca	ctgtccacga	tccagtacaa	cgtggtggat	ggcctgaggg	accgccggag	4380
cttccacggg	ccgtacacag	tgcaggccgg	gttgcctctg	aaccccatgg	gccgcacagg	4440
actgcgtggg	cgcgggagcc	tcagctgctt	cggaccacaac	cacacgctgt	accccatggg	4500
cacgcgggtg	aggcggaacg	aggatggagc	catctgcagg	aagagcataa	agaagatgct	4560
ggaagtgtctg	gtggtgaagc	tccctctctc	cgagcactgg	gccctgcctg	ggggctcccc	4620
ggagccaggg	gagatgctac	ctcgggaagct	gaagcggatc	ctccggcagg	agcactggcc	4680
gtcttttgaa	aacttgctga	agtgcggcat	ggaggtgtac	aaaggctaca	tggatgacct	4740
gaggaacacg	gacaatgcct	ggatcgagac	gggtgcgcgc	agcgtccact	tccaggacca	4800
gaatgacgtg	gagctgaaca	ggctgaactc	taacctgcac	gcctgcgact	cgggggcctc	4860
catccgatgg	caggtggtgg	acaggcgcat	cccactctat	gcgaaccaca	agaccctcct	4920
ccagaaggca	gccgctgagt	tcggggctca	ctactgactg	tgccctcagg	ctggggcggt	4980
caggtccata	gacgttcccc	ccagaaacca	gggttctctc	ctcctgagcc	tggccaggac	5040
tcaggctggt	cctgggcccct	gcacatgatg	gggtttgggtg	gacccagtgc	ccctcacggc	5100
tgccgcaagt	ctgctgcaga	tgacctcatg	aactggaagg	ggtcaagggtg	acccgggagg	5160
agagctcaag	acagggcaca	ggctactcag	agctgagggg	cccctgggac	ccttggccat	5220
caggcgaggg	gctgggcctg	tcagctggg	cccttgggca	gagtcactc	ccttctctggc	5280
tgtgtcacc	cagcagctc	atccaccatg	gaggtcattg	gcctgaggca	agttccccgg	5340
agagtgcggga	tccccctgtg	ccccctcagg	cctatgtctg	tgaggaaggg	gccctgccac	5400
tctccccaag	agggcctcca	tggttcgagg	tgctcaaca	tggagccttg	cctggcctgg	5460
gctaggggca	ctgtctgaac	tctgactgt	caggataaac	tccgtggggg	tacaggagcc	5520
cagacaaaagc	ccaggcctgt	caagagacgc	agaggccccc	tgccagggtt	ggccccaggg	5580
acctggggac	gaggtgcag	aagctctccc	tcctactcc	ctgggagcca	cgtgctggcc	5640
atgtggccag	ggacggcatg	agcaggaggc	ggggacgtgg	gggccttctg	gtttggtgtc	5700
aacagctcac	aggagcgtga	accatgaggg	ccctcaggag	gggaacgtgg	taaaaaccaa	5760
gacattaaat	ctgccatctc	aggcctggct	ggctcttctg	tgctttccac	aaataaagtt	5820
cctgacacgt	ccaggggccag	gggtgtgtg	acggctgcct	gaagttctcc	tcgatcccc	5880
ggtgagcttc	ctgcagcctg	tggatgtcct	gcagcccctc	agccctaccc	ccaagtttct	5940
cctctgaccc	atcagctccc	tgtcttcatt	ttcctaaacc	tgggtccag	catcgtcccc	6000
aagcccacca	ggccaggatg	caggcatcca	catgccctcc	tccttggtt	ccctgctgtg	6060
gtggtgccaa	tgtgccctgg	caccctgca	gaggtccgg	atggagcctg	gggtgcctg	6120
gccactgagc	actggccgag	gtgatgcccc	cccttcctg	gacaggcctc	tgtcttccac	6180
ctgacccaaa	gctctctagc	caccccttg	tccccagtat			6220

<210> 12

<211> 1503

	<400>												12			
Met 1	Glu	Pro	Ser	Ala 5	Leu	Arg	Lys	Ala	Gly 10	Ser	Glu	Gln	Glu 15	Gly 15		
Phe	Glu	Gly	Leu 20	Pro	Arg	Arg	Val 25	Thr	Asp	Leu	Gly	Met 30	Val 30	Asn		
Leu	Arg	Arg 35	Ser	Asn	Ser	Ser	Leu 40	Phe	Lys	Ser	Trp	Arg 45	Leu	Gln Cys		
Pro	Phe 50	Gly	Asn	Asn	Asp	Lys 55	Gln	Glu	Ser	Leu	Ser 60	Ser	Trp	Ile Pro		
Glu 65	Asn	Ile	Lys	Lys 70	Lys	Glu	Cys	Val	Tyr	Phe 75	Val	Glu	Ser	Ser Lys 80		
Leu	Ser	Asp	Ala	Gly 85	Lys	Val	Val	Cys	Gln 90	Cys	Gly	Tyr	Thr	His 95		
Gln	His	Leu	Glu 100	Glu	Ala	Thr	Lys	Pro 105	His	Thr	Phe	Gln	Gly 110	Thr Gln		
Trp	Asp	Pro 115	Lys	Lys	His	Val	Gln	Glu 120	Met	Pro	Thr	Asp 125	Ala	Phe Gly		
Asp	Ile 130	Val	Phe	Thr	Gly	Leu 135	Ser	Gln	Lys	Val	Lys 140	Lys	Tyr	Val Arg		
Val 145	Ser	Gln	Asp	Thr	Pro	Ser 150	Ser	Val	Ile	Tyr	His 155	Leu	Met	Thr Gln 160		
His	Trp	Gly	Leu	Asp 165	Val	Pro	Asn	Leu	Leu 170	Ile	Ser	Val	Thr	Gly 175		
Ala	Lys	Asn	Phe 180	Asn	Met	Lys	Pro	Arg 185	Leu	Lys	Ser	Ile	Phe 190	Arg Arg		
Gly	Leu	Val 195	Lys	Val	Ala	Gln	Thr 200	Thr	Gly	Ala	Trp	Ile 205	Ile	Thr Gly		
Gly	Ser 210	His	Thr	Gly	Val	Met 215	Lys	Gln	Val	Gly	Glu 220	Ala	Val	Arg Asp		
Phe 225	Ser	Leu	Ser	Ser	Ser 230	Tyr	Lys	Glu	Gly	Glu 235	Leu	Ile	Thr	Ile Gly 240		
Val	Ala	Thr	Trp	Gly 245	Thr	Val	His	Arg	Arg 250	Glu	Gly	Leu	Ile	His Pro 255		
Thr	Gly	Ser	Phe 260	Pro	Ala	Glu	Tyr	Ile 265	Leu	Asp	Glu	Asp	Gly 270	Gln Gly		
Asn	Leu	Thr 275	Cys	Leu	Asp	Ser	Asn 280	His	Ser	His	Phe	Ile 285	Leu	Val Asp		
Asp	Gly 290	Thr	His	Gly	Gln	Tyr 295	Gly	Val	Glu	Ile	Pro 300	Leu	Arg	Thr Arg		
Leu 305	Glu	Lys	Phe	Ile	Ser 310	Glu	Gln	Thr	Lys	Glu 315	Arg	Gly	Gly	Val Ala 320		
Ile	Lys	Ile	Pro	Ile 325	Val	Cys	Val	Val	Leu 330	Glu	Gly	Gly	Pro	Gly Thr 335		
Leu	His	Thr 340	Ile	Asp	Asn	Ala	Thr	Thr 345	Asn	Gly	Thr	Pro	Cys 350	Val Val		
Val	Glu	Gly 355	Ser	Gly	Arg	Val	Ala 360	Asp	Val	Ile	Ala	Gln 365	Val	Ala Asn		
Leu	Pro 370	Val	Ser	Asp	Ile	Thr 375	Ile	Ser	Leu	Ile	Gln 380	Gln	Lys	Leu Ser		
Val 385	Phe	Phe	Gln	Glu	Met 390	Phe	Glu	Thr	Phe	Thr 395	Glu	Ser	Arg	Ile Val 400		
Glu	Trp	Thr	Lys	Lys 405	Ile	Gln	Asp	Ile	Val 410	Arg	Arg	Arg	Gln	Leu Leu 415		
Thr	Val	Phe	Arg 420	Glu	Gly	Lys	Asp	Gly 425	Gln	Gln	Asp	Val	Asp 430	Val Ala		
Ile	Leu	Gln 435	Ala	Leu	Leu	Lys	Ala 440	Ser	Arg	Ser	Gln	Asp 445	His	Phe Gly		

-16-

His	Glu	Asn	Trp	Asp	His	Gln	Leu	Lys	Leu	Ala	Val	Ala	Trp	Asn	Arg
450					455					460					
Val	Asp	Ile	Ala	Arg	Ser	Glu	Ile	Phe	Met	Asp	Glu	Trp	Gln	Trp	Lys
465					470					475					480
Pro	Ser	Asp	Leu	His	Pro	Thr	Met	Thr	Ala	Ala	Leu	Ile	Ser	Asn	Lys
				485					490					495	
Pro	Glu	Phe	Val	Lys	Leu	Phe	Leu	Glu	Asn	Gly	Val	Gln	Leu	Lys	Glu
			500					505					510		
Phe	Val	Thr	Trp	Asp	Thr	Leu	Leu	Tyr	Leu	Tyr	Glu	Asn	Leu	Asp	Pro
		515				520					525				
Ser	Cys	Leu	Phe	His	Ser	Lys	Leu	Gln	Lys	Val	Leu	Val	Glu	Asp	Pro
530					535						540				
Glu	Arg	Pro	Ala	Cys	Ala	Pro	Ala	Ala	Pro	Arg	Leu	Gln	Met	His	His
545				550						555					560
Val	Ala	Gln	Val	Leu	Arg	Glu	Leu	Leu	Gly	Asp	Phe	Thr	Gln	Pro	Leu
				565					570					575	
Tyr	Pro	Arg	Pro	Arg	His	Asn	Asp	Arg	Leu	Arg	Leu	Leu	Leu	Pro	Val
			580					585					590		
Pro	His	Val	Lys	Leu	Asn	Val	Gln	Gly	Val	Ser	Leu	Arg	Ser	Leu	Tyr
		595				600					605				
Lys	Arg	Ser	Ser	Gly	His	Val	Thr	Phe	Thr	Met	Asp	Pro	Ile	Arg	Asp
610					615						620				
Leu	Leu	Ile	Trp	Ala	Ile	Val	Gln	Asn	Arg	Arg	Glu	Leu	Ala	Gly	Ile
625				630						635					640
Ile	Trp	Ala	Gln	Ser	Gln	Asp	Cys	Ile	Ala	Ala	Ala	Leu	Ala	Cys	Ser
				645					650					655	
Lys	Ile	Leu	Lys	Glu	Leu	Ser	Lys	Glu	Glu	Glu	Asp	Thr	Asp	Ser	Ser
			660					665					670		
Glu	Glu	Met	Leu	Ala	Leu	Ala	Glu	Glu	Tyr	Glu	His	Arg	Ala	Ile	Gly
		675					680					685			
Val	Phe	Thr	Glu	Cys	Tyr	Arg	Lys	Asp	Glu	Glu	Arg	Ala	Gln	Lys	Leu
		690				695					700				
Leu	Thr	Arg	Val	Ser	Glu	Ala	Trp	Gly	Lys	Thr	Thr	Cys	Leu	Gln	Leu
705				710						715					720
Ala	Leu	Glu	Ala	Lys	Asp	Met	Lys	Phe	Val	Ser	His	Gly	Gly	Ile	Gln
				725					730					735	
Ala	Phe	Leu	Thr	Lys	Val	Trp	Trp	Gly	Gln	Leu	Ser	Val	Asp	Asn	Gly
			740					745					750		
Leu	Trp	Arg	Val	Thr	Leu	Cys	Met	Leu	Ala	Phe	Pro	Leu	Leu	Leu	Thr
		755					760					765			
Gly	Leu	Ile	Ser	Phe	Arg	Glu	Lys	Arg	Leu	Gln	Asp	Val	Gly	Thr	Pro
770					775						780				
Ala	Ala	Arg	Ala	Arg	Ala	Phe	Phe	Thr	Ala	Pro	Val	Val	Val	Phe	His
785					790					795					800
Leu	Asn	Ile	Leu	Ser	Tyr	Phe	Ala	Phe	Leu	Cys	Leu	Phe	Ala	Tyr	Val
				805					810					815	
Leu	Met	Val	Asp	Phe	Gln	Pro	Val	Pro	Ser	Trp	Cys	Glu	Cys	Ala	Ile
			820					825					830		
Tyr	Leu	Trp	Leu	Phe	Ser	Leu	Val	Cys	Glu	Glu	Met	Arg	Gln	Leu	Phe
		835				840						845			
Tyr	Asp	Pro	Asp	Glu	Cys	Gly	Leu	Met	Lys	Lys	Ala	Ala	Leu	Tyr	Phe
850						855					860				
Ser	Asp	Phe	Trp	Asn	Lys	Leu	Asp	Val	Gly	Ala	Ile	Leu	Leu	Phe	Val
865				870						875					880
Ala	Gly	Leu	Thr	Cys	Arg	Leu	Ile	Pro	Ala	Thr	Leu	Tyr	Pro	Gly	Arg
				885					890					895	
Val	Ile	Leu	Ser	Leu	Asp	Phe	Ile	Leu	Phe	Cys	Leu	Arg	Leu	Met	His
			900					905					910		
Ile	Phe	Thr	Ile	Ser	Lys	Thr	Leu	Gly	Pro	Lys	Ile	Ile	Ile	Val	Lys
		915					920						925		

-17-

Arg	Met	Met	Lys	Asp	Val	Phe	Phe	Phe	Leu	Phe	Leu	Leu	Ala	Val	Trp
930						935					940				
Val	Val	Ser	Phe	Gly	Val	Ala	Lys	Gln	Ala	Ile	Leu	Ile	His	Asn	Glu
945					950					955					960
Arg	Arg	Val	Asp	Trp	Leu	Phe	Arg	Gly	Ala	Val	Tyr	His	Ser	Tyr	Leu
			965						970						975
Thr	Ile	Phe	Gly	Gln	Ile	Pro	Gly	Tyr	Ile	Asp	Gly	Val	Asn	Phe	Asn
			980					985					990		
Pro	Glu	His	Cys	Ser	Pro	Asn	Gly	Thr	Asp	Pro	Tyr	Lys	Pro	Lys	Cys
		995					1000					1005			
Pro	Glu	Ser	Asp	Ala	Thr	Gln	Gln	Arg	Pro	Ala	Phe	Pro	Glu	Trp	Leu
		1010				1015					1020				
Thr	Val	Leu	Leu	Leu	Cys	Leu	Tyr	Leu	Leu	Phe	Thr	Asn	Ile	Leu	Leu
1025					1030					1035					104
Leu	Asn	Leu	Leu	Ile	Ala	Met	Phe	Asn	Tyr	Thr	Phe	Gln	Gln	Val	Gln
				1045					1050						1055
Glu	His	Thr	Asp	Gln	Ile	Trp	Lys	Phe	Gln	Arg	His	Asp	Leu	Ile	Glu
			1060					1065					1070		
Glu	Tyr	His	Gly	Arg	Pro	Ala	Ala	Pro	Pro	Pro	Phe	Ile	Leu	Leu	Ser
		1075				1080						1085			
His	Leu	Gln	Leu	Phe	Ile	Lys	Arg	Val	Val	Leu	Lys	Thr	Pro	Ala	Lys
		1090				1095					1100				
Arg	His	Lys	Gln	Leu	Lys	Asn	Lys	Leu	Glu	Lys	Asn	Glu	Glu	Ala	Ala
1105					1110					1115					112
Leu	Leu	Ser	Trp	Glu	Ile	Tyr	Leu	Lys	Glu	Asn	Tyr	Leu	Gln	Asn	Arg
			1125						1130						1135
Gln	Phe	Gln	Gln	Lys	Gln	Arg	Pro	Glu	Gln	Lys	Ile	Glu	Asp	Ile	Ser
			1140					1145					1150		
Asn	Lys	Val	Asp	Ala	Met	Val	Asp	Leu	Leu	Asp	Leu	Asp	Pro	Leu	Lys
		1155					1160					1165			
Arg	Ser	Gly	Ser	Met	Glu	Gln	Arg	Leu	Ala	Ser	Leu	Glu	Glu	Gln	Val
		1170				1175					1180				
Ala	Gln	Thr	Ala	Arg	Ala	Leu	His	Trp	Ile	Val	Arg	Thr	Leu	Arg	Ala
1185					1190					1195					120
Ser	Gly	Phe	Ser	Ser	Glu	Ala	Asp	Val	Pro	Thr	Leu	Ala	Ser	Gln	Lys
			1205						1210					1215	
Ala	Ala	Glu	Glu	Pro	Asp	Ala	Glu	Pro	Gly	Gly	Arg	Lys	Lys	Thr	Glu
			1220					1225					1230		
Glu	Pro	Gly	Asp	Ser	Tyr	His	Val	Asn	Ala	Arg	His	Leu	Leu	Tyr	Pro
		1235					1240					1245			
Asn	Cys	Pro	Val	Thr	Arg	Phe	Pro	Val	Pro	Asn	Glu	Lys	Val	Pro	Trp
		1250				1255					1260				
Glu	Thr	Glu	Phe	Leu	Ile	Tyr	Asp	Pro	Pro	Phe	Tyr	Thr	Ala	Glu	Arg
1265					1270					1275					128
Lys	Asp	Ala	Ala	Ala	Met	Asp	Pro	Met	Gly	Asp	Thr	Leu	Glu	Pro	Leu
			1285						1290					1295	
Ser	Thr	Ile	Gln	Tyr	Asn	Val	Val	Asp	Gly	Leu	Arg	Asp	Arg	Arg	Ser
			1300					1305					1310		
Phe	His	Gly	Pro	Tyr	Thr	Val	Gln	Ala	Gly	Leu	Pro	Leu	Asn	Pro	Met
		1315					1320					1325			
Gly	Arg	Thr	Gly	Leu	Arg	Gly	Arg	Gly	Ser	Leu	Ser	Cys	Phe	Gly	Pro
		1330				1335					1340				
Asn	His	Thr	Leu	Tyr	Pro	Met	Val	Thr	Arg	Trp	Arg	Arg	Asn	Glu	Asp
1345					1350					1355					136
Gly	Ala	Ile	Cys	Arg	Lys	Ser	Ile	Lys	Lys	Met	Leu	Glu	Val	Leu	Val
			1365					1370						1375	
Val	Lys	Leu	Pro	Leu	Ser	Glu	His	Trp	Ala	Leu	Pro	Gly	Gly	Ser	Arg
			1380					1385					1390		
Glu	Pro	Gly	Glu	Met	Leu	Pro	Arg	Lys	Leu	Lys	Arg	Ile	Leu	Arg	Gln
		1395					1400						1405		

-18-

Glu His Trp Pro Ser Phe Glu Asn Leu Leu Lys Cys Gly Met Glu Val
 1410 1415 1420
 Tyr Lys Gly Tyr Met Asp Asp Pro Arg Asn Thr Asp Asn Ala Trp Ile
 1425 1430 1435 144
 Glu Thr Val Ala Val Ser Val His Phe Gln Asp Gln Asn Asp Val Glu
 1445 1450 1455
 Leu Asn Arg Leu Asn Ser Asn Leu His Ala Cys Asp Ser Gly Ala Ser
 1460 1465 1470
 Ile Arg Trp Gln Val Val Asp Arg Arg Ile Pro Leu Tyr Ala Asn His
 1475 1480 1485
 Lys Thr Leu Leu Gln Lys Ala Ala Ala Glu Phe Gly Ala His Tyr
 1490 1495 1500

<210> 13
 <211> 1816
 <212> PRT
 <213> C. Elegans

<400> 13
 Met Ile Thr Asp Lys Asn Leu Phe Ser Arg Leu Leu Ile Lys Lys Asn
 1 5 10 15
 Pro Ile Arg Met His Ser Pro Ser Phe Ser Phe Ser Leu Ile Thr Ser
 20 25 30
 Leu Phe Phe Thr Gln Phe Phe Met Phe Gln Leu Ser Ser Met Ala Tyr
 35 40 45
 Phe Phe Leu Thr Leu Ile Ala Gly Val Thr His Phe Tyr Phe Pro Glu
 50 55 60
 Lys Leu Leu Gly Lys Ser Glu Asn Leu Asp His Arg Tyr Gln Ser Ser
 65 70 75 80
 Glu Gln Lys Val Leu Ile Glu Trp Thr Glu Asn Lys Ala Val Ala Glu
 85 90 95
 Ser Leu Arg Ala Asn Ser Val Thr Val Glu Glu Asn Glu Ser Glu Arg
 100 105 110
 Glu Thr Glu Thr Gln Thr Lys Arg Arg Lys Lys Gln Arg Ser Thr
 115 120 125
 Ser Ser Asp Lys Ala Pro Leu Asn Ser Ala Pro Arg His Val Gln Lys
 130 135 140
 Phe Asp Trp Lys Asp Met Leu His Leu Ala Asp Ile Ser Gly Arg Lys
 145 150 155 160
 Arg Gly Asn Ser Thr Thr Ser His Ser Gly His Ala Thr Arg Ala Gly
 165 170 175
 Ser Leu Lys Gly Lys Asn Trp Ile Glu Cys Arg Leu Lys Met Arg Gln
 180 185 190
 Cys Ser Tyr Phe Val Pro Ser Gln Arg Phe Ser Glu Arg Cys Gly Cys
 195 200 205
 Gly Lys Glu Arg Ser Lys His Thr Glu Glu Val Leu Glu Arg Ser Gln
 210 215 220
 Asn Lys Asn His Pro Leu Asn His Leu Thr Leu Pro Gly Ile His Glu
 225 230 235 240
 Val Asp Thr Thr Asp Ala Asp Ala Asp Asp Asn Glu Val Asn Leu Thr
 245 250 255
 Pro Gly Arg Trp Ser Ile Gln Ser His Thr Glu Ile Val Pro Thr Asp
 260 265 270
 Ala Tyr Gly Asn Ile Val Phe Glu Gly Thr Ala His His Ala Gln Tyr
 275 280 285
 Ala Arg Ile Ser Phe Asp Ser Asp Pro Arg Asp Ile Val His Leu Met
 290 295 300
 Met Lys Val Trp Lys Leu Lys Pro Pro Lys Leu Ile Ile Thr Ile Asn
 305 310 315 320
 Gly Gly Leu Thr Lys Phe Asp Leu Gln Pro Lys Leu Ala Arg Thr Phe

				325					330					335		
Arg	Lys	Gly	Ile	Met	Lys	Ile	Ala	Lys	Ser	Thr	Asp	Ala	Trp	Ile	Ile	
			340					345					350			
Thr	Ser	Gly	Leu	Asp	Glu	Gly	Val	Val	Lys	His	Leu	Asp	Ser	Ala	Leu	
		355					360					365				
His	Ala	Leu	Glu	Phe	Trp	Ser	Phe	Gly	Leu	Phe	Trp	Val	Ile	Gln	Leu	
		370				375					380					
Asp	Val	Leu	Leu	Ala	His	Ser	Met	Phe	Ile	Pro	Arg	Gly	Ser	Leu	Phe	
385					390					395					400	
Asp	His	Gly	Asn	His	Thr	Ser	Lys	Asn	His	Val	Val	Ala	Ile	Gly	Ile	
			405					410						415		
Ala	Ser	Trp	Gly	Met	Leu	Lys	Gln	Arg	Ser	Arg	Phe	Val	Gly	Lys	Asp	
			420					425					430			
Ser	Thr	Val	Thr	Tyr	Ala	Thr	Asn	Val	Phe	Asn	Asn	Thr	Arg	Leu	Lys	
		435					440					445				
Glu	Leu	Asn	Asp	Asn	His	Ser	Tyr	Phe	Leu	Phe	Ser	Asp	Asn	Gly	Thr	
		450				455					460					
Val	Asn	Arg	Tyr	Gly	Ala	Glu	Ile	Ile	Met	Arg	Lys	Arg	Leu	Glu	Ala	
465					470					475					480	
Tyr	Leu	Ala	Gln	Gly	Asp	Lys	Lys	Arg	Ser	Ala	Ile	Pro	Leu	Val	Cys	
			485					490					495			
Val	Val	Leu	Glu	Gly	Ala	Phe	Thr	Ile	Lys	Met	Val	His	Asp	Tyr		
			500				505					510				
Val	Thr	Thr	Ile	Pro	Arg	Ile	Pro	Val	Ile	Val	Cys	Asp	Gly	Ser	Gly	
		515					520					525				
Arg	Ala	Ala	Asp	Ile	Leu	Ala	Phe	Ala	His	Gln	Ala	Val	Ser	Gln	Asn	
		530				535					540					
Gly	Phe	Leu	Ser	Asp	Asn	Ile	Arg	Asn	Gln	Leu	Val	Asn	Ile	Val	Arg	
545					550					555					560	
Arg	Ile	Phe	Gly	Tyr	Asp	Pro	Lys	Thr	Ala	Gln	Lys	Leu	Ile	Lys	Gln	
			565						570					575		
Ile	Val	Glu	Cys	Ser	Thr	Asn	Lys	Ser	Leu	Met	Thr	Ile	Phe	Arg	Leu	
			580					585					590			
Gly	Glu	Ser	Ser	Arg	Glu	Asp	Leu	Asp	His	Val	Ile	Met	Ser	Cys	Leu	
		595					600					605				
Leu	Lys	Gly	Gln	Asn	Leu	Ser	Pro	Pro	Glu	Gln	Leu	Gln	Leu	Ala	Leu	
		610				615					620					
Ala	Trp	Asn	Arg	Ala	Asp	Ile	Ala	Arg	Thr	Glu	Ile	Phe	Ala	Asn	Gly	
625					630					635					640	
Thr	Glu	Trp	Thr	Thr	Gln	Asp	Leu	His	Asn	Ala	Met	Ile	Glu	Ala	Leu	
			645						650					655		
Ser	Asn	Asp	Arg	Ile	Asp	Phe	Val	His	Leu	Leu	Leu	Glu	Asn	Gly	Val	
			660													

-20-

				805					810				815		
Asn	Met	Asp	Phe	Thr	Phe	Arg	Tyr	Pro	Tyr	Ser	Asp	Leu	Met	Ile	Trp
			820					825					830		
Ala	Val	Leu	Thr	Lys	Arg	Gln	Lys	Met	Ala	Lys	Leu	Met	Trp	Thr	His
		835					840					845			
Gly	Glu	Glu	Gly	Met	Ala	Lys	Ala	Leu	Val	Ala	Ser	Arg	Leu	Tyr	Val
	850					855					860				
Ser	Leu	Ala	Lys	Thr	Ala	Ser	Leu	Ala	Thr	Gly	Glu	Ile	Gly	Met	Ser
865				870						875					880
Gln	Asp	Phe	Thr	Glu	Phe	Ser	Asp	Glu	Phe	Ser	Glu	Leu	Ala	Val	Glu
			885					890						895	
Val	Leu	Glu	Tyr	Cys	Thr	Lys	His	Gly	Arg	Asp	Gln	Thr	Leu	Arg	Leu
		900						905					910		
Leu	Thr	Cys	Glu	Leu	Ala	Asn	Trp	Gly	Asp	Glu	Thr	Cys	Leu	Ser	Leu
		915					920					925			
Ala	Ala	Asn	Asn	Gly	His	Arg	Lys	Phe	Leu	Ala	His	Pro	Cys	Cys	Gln
	930					935					940				
Met	Leu	Leu	Ser	Asp	Leu	Trp	Gln	Gly	Gly	Leu	Leu	Met	Lys	Asn	Asn
945				950						955					960
Gln	Asn	Ser	Lys	Val	Leu	Thr	Cys	Leu	Ala	Ala	Pro	Pro	Leu	Ile	Phe
			965					970						975	
Leu	Leu	Gly	Phe	Lys	Thr	Lys	Glu	Gln	Leu	Met	Leu	Gln	Pro	Lys	Thr
		980					985					990			
Ala	Ala	Glu	His	Asp	Glu	Glu	Met	Ser	Asp	Ser	Glu	Met	Asn	Ser	Ala
	995						1000					1005			
Glu	Asp	Thr	Asp	Thr	Ser	Ser	Asp	Ser	Ser	Ser	Asp	Ser	Asp	Asp	Ser
	1010					1015					1020				
Asp	Glu	Glu	Asp	Ala	Lys	Leu	Arg	Ala	Gln	Ser	Leu	Ser	Ala	Asp	Gln
1025				1030					1035						104
Pro	Leu	Ser	Ile	His	Arg	Leu	Val	Arg	Asp	Lys	Leu	Asn	Phe	Ser	Glu
			1045					1050						1055	
Lys	Lys	Lys	Pro	Asp	Met	Gly	Ile	Ser	Arg	Ile	Val	Val	Ala	Pro	Pro
		1060				1065						1070			
Ile	Val	Thr	Gly	Arg	Asn	Arg	Ala	Arg	Thr	Met	Ser	Ile	Lys	Lys	Ser
	1075					1080					1085				
Lys	Lys	Asn	Val	Ile	Lys	Pro	Pro	Ala	Cys	Leu	Lys	Ile	Glu	Thr	Ser
	1090					1095					1100				
Asp	Asp	Asp	Glu	Gln	Glu	Lys	Lys	Ala	Thr	Glu	Met	Cys	Lys	Ser	
1105				1110						1115					112
Thr	Phe	Phe	Asp	Phe	Phe	Phe	Asp	Phe	Pro	Tyr	Ile	Asn	Arg	Thr	Gly
			1125						1130					1135	
Lys	Arg	Gly	Ser	Val	Ala	Val	Ala	Met	Asn	His	Asp	Asp	Met	Tyr	Ile
		1140						1145					1150		
Asp	Pro	Ser	Glu	Glu	Leu	Asp	Thr	Gln	Thr	Arg	Gln	Lys	Ser	Ser	Arg
	1155					1160					1165				
Glu	Phe	Ser	Ser	Ser	Arg	Asn	Val	Thr	Val	Gln	Val	Tyr	Thr	Gln	Arg
	1170					1175					1180				
Pro	Leu	Ser	Trp	Lys	Lys	Lys	Ile	Met	Glu	Phe	Tyr	Lys	Ala	Pro	Ile
1185				1190						1195					120
Thr	Thr	Tyr	Trp	Leu	Trp	Phe	Phe	Ala	Phe	Ile	Trp	Phe	Leu	Ile	Leu
			1205					1210						1215	
Leu	Thr	Tyr	Asn	Leu	Leu	Val	Lys	Thr	Gln	Arg	Ile	Ala	Ser	Trp	Ser
		1220					1225					1230			
Glu	Trp	Tyr	Val	Phe	Ala	Tyr	Ile	Phe	Val	Trp	Thr	Leu	Glu	Ile	Gly
	1235					1240					1245				
Arg	Lys	Val	Val	Ser	Thr	Ile	Met	Met	Asp	Thr	Ser	Lys	Pro	Val	Leu
	1250					1255					1260				
Lys	Gln	Leu	Arg	Val	Phe	Phe	Phe	Gln	Tyr	Arg	Asn	Gly	Leu	Leu	Ala
1265				1270						1275					128
Phe	Gly	Leu	Leu	Thr	Tyr	Leu	Ile	Ala	Tyr	Phe	Ile	Arg	Leu	Ser	Pro

-21-

				1285						1290					1295
Thr	Thr	Lys	Thr	Leu	Gly	Arg	Ile	Leu	Ile	Ile	Cys	Asn	Ser	Val	Ile
			1300					1305						1310	
Trp	Ser	Leu	Lys	Leu	Val	Asp	Tyr	Leu	Ser	Val	Gln	Gln	Gly	Leu	Gly
		1315					1320						1325		
Pro	Tyr	Ile	Asn	Ile	Val	Ala	Glu	Met	Ile	Pro	Thr	Met	Ile	Pro	Leu
	1330				1335					1340					
Cys	Val	Leu	Val	Phe	Ile	Thr	Leu	Tyr	Ala	Phe	Gly	Leu	Leu	Arg	Gln
1345					1350					1355					136
Ser	Ile	Thr	Tyr	Pro	Tyr	Glu	Asp	Trp	His	Trp	Ile	Leu	Val	Arg	Asn
				1365					1370					1375	
Ile	Phe	Leu	Gln	Pro	Tyr	Phe	Met	Leu	Tyr	Gly	Glu	Val	Tyr	Ala	Ala
		1380					1385						1390		
Glu	Ile	Asp	Thr	Cys	Gly	Asp	Glu	Ile	Trp	Gln	Thr	His	Glu	Asp	Glu
	1395						1400					1405			
Asn	Ile	Pro	Ile	Ser	Met	Leu	Asn	Val	Thr	His	Glu	Thr	Cys	Val	Pro
	1410				1415						1420				
Gly	Tyr	Trp	Ile	Ala	Pro	Val	Gly	Leu	Thr	Val	Phe	Met	Leu	Ala	Thr
1425					1430					1435					144
Asn	Val	Leu	Leu	Met	Asn	Val	Met	Val	Ala	Gly	Cys	Thr	Tyr	Ile	Phe
				1445					1450					1455	
Glu	Lys	His	Ile	Gln	Ser	Thr	Arg	Glu	Ile	Phe	Leu	Phe	Glu	Arg	Tyr
		1460						1465					1470		
Gly	Gln	Val	Met	Glu	Tyr	Glu	Ser	Thr	Pro	Trp	Leu	Pro	Pro	Pro	Phe
	1475						1480					1485			
Thr	Ile	Ile	Tyr	His	Val	Ile	Trp	Leu	Phe	Lys	Leu	Ile	Lys	Ser	Ser
	1490					1495						1500			
Ser	Arg	Met	Phe	Glu	Arg	Lys	Asn	Leu	Phe	Asp	Gln	Ser	Leu	Lys	Leu
1505					1510					1515					152
Phe	Leu	Ser	Pro	Asp	Glu	Met	Glu	Lys	Val	His	Thr	Phe	Glu	Glu	Glu
				1525					1530					1535	
Ser	Val	Glu	Asp	Met	Lys	Arg	Glu	Thr	Glu	Lys	Lys	Asn	Leu	Ser	Ser
		1540						1545					1550		
Asn	Asp	Glu	Arg	Ile	His	Arg	Thr	Ala	Glu	Arg	Thr	Asp	Ala	Ile	Leu
	1555						1560					1565			
Asn	Arg	Val	Ser	His	Leu	Thr	Gln	Leu	Glu	Phe	Thr	Leu	Lys	Glu	Glu
	1570					1575					1580				
Ile	Arg	Glu	Leu	Glu	His	Lys	Met	Lys	Asn	Met	Asp	Ser	Arg	His	Lys
1585					1590					1595					160
Glu	Gln	Met	Asn	Leu	Met	Leu	Asp	Met	Asn	Lys	Lys	Leu	Gly	Lys	Phe
			1605						1610					1615	
Ile	Ser	Gly	Lys	Tyr	Lys	Arg	Gly	Ser	Phe	Gly	Gly	Ser	Gly	Ser	Asp
		1620						1625					1630		
Gly	Gly	Gly	Gly	Ser	Ser	Asp	Asn	Ser	Lys	Leu	Glu	Pro	Asn	Asn	Ser
	1635						1640					1645			
Val	Pro	Met	Ile	Thr	Val	Asp	Gly	Pro	Ser	Pro	Ile	Gly	Ser	Arg	Arg
	1650					1655					1660				
Thr	Ser	Gly	Gln	Tyr	Leu	Lys	Arg	Asp	Ser	Leu	Gln	Ala	Lys	Lys	Lys
1665					1670					1675					168
Ile	Thr	Glu	Asn	Arg	Arg	Ser	Ser	Leu	Glu	Gln	Pro	Lys	Ile	Pro	Ser
			1685					1690						1695	
Ile	Gln	Phe	Asn	Leu	Met	Glu	Asp	Gln	Asp	Glu	Ser	Ala	Ala	Glu	Ser
	1700						1705					1710			
Ala	Thr	Glu	Glu	Val	Ser	Ile	Ser	Ile	Pro	Val	Pro	Gln	Met	Arg	Val
	1715						1720					1725			
Arg	Gln	Val	Thr	Glu	Ser	Asp	Lys	Ser	Asp	Leu	Ser	Glu	Asp	Asp	Leu
	1730					1735					1740				
Ile	Thr	Arg	Glu	Asp	Ala	Pro	Pro	Thr	Ser	Ile	Asn	Leu	Pro	Arg	Gly
1745					1750					1755					176
Pro	Arg	Arg	His	Ala	Leu	Tyr	Ser	Thr	Ile	Ala	Asp	Ala	Ile	Glu	Thr

```
<210> 14
<211> 1387
<212> PRT
<213> C. Elegans
```

<400> 14															
Met 1	Arg	Lys	Ser 5	Arg	Arg	Val	Arg	Lys	Leu 10	Val	Arg	His	Ala	Ser 15	Leu
Ile	Glu	Asn	Ile 20	Arg	His	Arg	Thr	Ser 25	Ser	Phe	Leu	Arg	Leu 30	Leu	Asn
Ala	Pro	Arg 35	Asn	Ser	Met	Cys	Asn 40	Ala	Asn	Thr	Val	His 45	Ser	Ile	Ser
Ser	Phe	Arg 50	Ser	Asp	His	Leu 55	Ser	Arg	Lys	Ser	Thr 60	His	Lys	Phe	Leu
Asp 65	Asn	Pro	Asn	Leu 70	Phe	Ala	Ile	Glu	Leu 75	Thr	Glu	Lys	Leu	Ser	Pro 80
Pro	Trp	Ile	Glu 85	Asn	Thr	Phe	Glu	Lys 90	Arg	Glu	Cys	Ile	Arg	Phe 95	Ala
Ala	Leu	Pro	Lys 100	Asp	Pro	Glu	Arg	Cys 105	Gly	Cys	Gly	Arg	Pro 110	Leu	Ser
Ala	His	Thr 115	Pro	Ala	Ser	Thr	Phe 120	Ser	Thr	Leu	Pro 125	Val	His	Leu	
Leu	Glu	Lys 130	Glu	Gln	Gln	Thr 135	Trp	Thr	Ile	Ala	Asn 140	Asn	Thr	Gln	Thr
Ser 145	Thr	Thr	Asp	Ala 150	Phe	Gly	Thr	Ile	Val 155	Phe	Gln	Gly	Gly	Ala	His 160
Ala	His	Lys	Ala 165	Gln	Tyr	Val	Arg	Leu 170	Ser	Tyr	Asp	Ser	Glu	Pro 175	Leu
Asp	Val	Met	Tyr 180	Leu	Met	Glu	Lys 185	Val	Trp	Gly	Leu	Glu	Ala 190	Pro	Arg
Leu	Val	Ile 195	Thr	Val	His	Gly	Gly 200	Met	Ser	Asn	Phe	Glu	Leu 205	Glu	Glu
Arg	Leu	Gly 210	Arg	Leu	Phe 215	Arg	Lys	Gly	Met	Leu	Lys 220	Ala	Ala	Gln	Thr
Thr 225	Gly	Ala	Trp	Ile 230	Ile	Thr	Ser	Gly	Leu 235	Asp	Ser	Gly	Val	Val	Arg 240
His	Val	Ala	Lys 245	Ala	Leu	Asp	Glu	Ala 250	Gly	Ile	Ser	Ala	Arg	Met	Arg 255
Ser	Gln	Ile	Val 260	Thr	Ile	Gly	Ile 265	Ala	Pro	Trp	Gly	Val	Ile 270	Lys	Arg
Lys	Glu	Arg 275	Leu	Ile	Arg	Gln	Asn 280	Glu	His	Val	Tyr	Tyr 285	Asp	Val	His
Ser	Leu	Ser 290	Val	Asn	Ala	Asn 295	Val	Gly	Ile	Leu	Asn 300	Asp	Arg	His	Ser
Tyr 305	Phe	Leu	Leu	Ala 310	Asp	Asn	Gly	Thr	Val 315	Gly	Arg	Phe	Gly	Ala	Asp 320
Leu	His	Leu	Arg 325	Gln	Asn	Leu	Glu	Asn 330	His	Ile	Ala	Thr	Phe	Gly 335	Cys
Asn	Gly	Arg	Lys 340	Val	Pro	Val	Val	Cys 345	Thr	Leu	Leu	Glu	Gly 350	Gly	Ile
Ser	Ser	Ile 355	Asn	Ala	Ile	His	Asp 360	Tyr	Val	Thr	Met	Lys 365	Pro	Asp	Ile

-23-

Pro	Ala	Ile	Val	Cys	Asp	Gly	Ser	Gly	Arg	Ala	Ala	Asp	Ile	Ile	Ser
370						375					380				
Phe	Ala	Ala	Arg	Tyr	Ile	Asn	Ser	Asp	Gly	Thr	Phe	Ala	Ala	Glu	Val
385					390					395					400
Gly	Glu	Lys	Leu	Arg	Asn	Leu	Ile	Lys	Met	Val	Phe	Pro	Glu	Thr	Asp
				405					410					415	
Gln	Glu	Glu	Met	Phe	Arg	Lys	Ile	Thr	Glu	Cys	Val	Ile	Arg	Asp	Asp
			420						425					430	
Leu	Leu	Arg	Ile	Phe	Arg	Tyr	Gly	Gln	Glu	Glu	Glu	Glu	Asp	Val	Asp
		435					440						445		
Phe	Val	Ile	Leu	Ser	Thr	Val	Leu	Gln	Lys	Gln	Asn	Leu	Pro	Pro	Asp
450						455					460				
Glu	Gln	Leu	Ala	Leu	Thr	Leu	Ser	Trp	Asn	Arg	Val	Asp	Leu	Ala	Lys
465					470					475					480
Ser	Cys	Leu	Phe	Ser	Asn	Gly	Arg	Lys	Trp	Ser	Ser	Asp	Val	Leu	Glu
				485					490					495	
Lys	Ala	Met	Asn	Asp	Ala	Leu	Tyr	Trp	Asp	Arg	Val	Asp	Phe	Val	Glu
			500					505					510		
Cys	Leu	Leu	Glu	Asn	Gly	Val	Ser	Met	Lys	Asn	Phe	Leu	Ser	Ile	Asn
			515				520						525		
Arg	Leu	Glu	Asn	Leu	Tyr	Asn	Met	Asp	Asp	Ile	Asn	Ser	Ala	His	Ser
					535							540			
Val	Arg	Asn	Trp	Met	Glu	Asn	Phe	Asp	Ser	Met	Asp	Pro	His	Thr	Tyr
545					550					555					560
Leu	Thr	Ile	Pro	Met	Ile	Gly	Gln	Val	Val	Glu	Lys	Leu	Met	Gly	Asn
				565					570					575	
Ala	Phe	Gln	Leu	Tyr	Tyr	Thr	Ser	Arg	Ser	Phe	Lys	Gly	Lys	Tyr	Asp
			580					585					590		
Arg	Tyr	Lys	Arg	Ile	Asn	Gln	Ser	Ser	Tyr	Phe	His	Arg	Lys	Arg	Lys
		595					600						605		
Ile	Val	Gln	Lys	Glu	Leu	Phe	Lys	Lys	Lys	Ser	Asp	Asp	Gln	Ile	Asn
610					615						620				
Asp	Asn	Glu	Glu	Glu	Asp	Phe	Ser	Phe	Ala	Tyr	Pro	Phe	Asn	Asp	Leu
625					630					635					640
Leu	Ile	Trp	Ala	Val	Leu	Thr	Ser	Arg	His	Gly	Met	Ala	Glu	Cys	Met
				645					650					655	
Trp	Val	His	Gly	Glu	Asp	Ala	Met	Ala	Lys	Cys	Leu	Leu	Ala	Ile	Arg
			660					665					670		
Leu	Tyr	Lys	Ala	Thr	Ala	Lys	Ile	Ala	Glu	Asp	Glu	Tyr	Leu	Asp	Val
		675					680					685			
Glu	Glu	Ala	Lys	Arg	Leu	Phe	Asp	Asn	Ala	Val	Lys	Cys	Arg	Glu	Asp
690					695						700				
Ala	Ile	Glu	Leu	Leu	Asp	Gln	Cys	Tyr	Arg	Ala	Asp	His	Asp	Arg	Thr
705					710					715					720
Leu	Arg	Leu	Leu	Arg	Met	Glu	Leu	Pro	His	Trp	Gly	Asn	Asn	Asn	Cys
				725					730					735	
Leu	Ser	Leu	Ala	Val	Leu	Ala	Asn	Thr	Lys	Thr	Phe	Leu	Ala	His	Pro
			740					745					750		
Cys	Cys	Gln	Ile	Leu	Leu	Ala	Glu	Leu	Trp	His	Gly	Ser	Leu	Lys	Val
		755					760					765			
Arg	Ser	Gly	Ser	Asn	Val	Arg	Val	Leu	Thr	Ala	Leu	Ile	Cys	Pro	Pro
770						775						780			
Ala	Ile	Leu	Phe	Met	Ala	Tyr	Lys	Pro	Lys	His	Ser	Lys	Thr	Ala	Arg
785					790					795					800
Leu	Leu	Ser	Glu	Glu	Thr	Pro	Glu	Gln	Leu	Pro	Tyr	Pro	Arg	Glu	Ser
				805					810					815	
Ile	Thr	Ser	Thr	Thr	Ser	Asn	Arg	Tyr	Arg	Tyr	Ser	Lys	Gly	Pro	Glu
			820					825					830		
Glu	Gln	Lys	Glu	Thr	Leu	Leu	Glu	Lys	Gly	Ser	Tyr	Thr	Lys	Lys	Val
		835					840						845		

-24-

Thr	Ile	Ile	Ser	Ser	Arg	Lys	Asn	Ser	Gly	Val	Ala	Ser	Val	Tyr	Gly
850						855					860				
Ser	Ala	Ser	Ser	Met	Met	Phe	Lys	Arg	Glu	Pro	Gln	Leu	Asn	Lys	Phe
865				870						875					880
Glu	Arg	Phe	Arg	Ala	Phe	Tyr	Ser	Ser	Pro	Ile	Thr	Lys	Phe	Trp	Ser
				885					890					895	
Trp	Cys	Ile	Ala	Phe	Leu	Ile	Phe	Leu	Thr	Thr	Gln	Thr	Cys	Ile	Leu
			900					905					910		
Leu	Leu	Glu	Thr	Ser	Leu	Lys	Pro	Ser	Lys	Tyr	Glu	Trp	Ile	Thr	Phe
		915					920					925			
Ile	Tyr	Thr	Val	Thr	Leu	Ser	Val	Glu	His	Ile	Arg	Lys	Leu	Met	Thr
		930				935					940				
Ser	Glu	Gly	Ser	Arg	Ile	Asn	Glu	Lys	Val	Lys	Val	Phe	Tyr	Ala	Lys
945					950					955					960
Trp	Tyr	Asn	Ile	Trp	Thr	Ser	Ala	Ala	Leu	Leu	Phe	Phe	Leu	Val	Gly
				965					970					975	
Tyr	Gly	Phe	Arg	Leu	Val	Pro	Met	Tyr	Arg	His	Ser	Trp	Gly	Arg	Val
			980					985					990		
Leu	Leu	Ser	Phe	Ser	Asn	Val	Leu	Phe	Tyr	Met	Lys	Ile	Phe	Glu	Tyr
		995					1000					1005			
Leu	Ser	Val	His	Pro	Leu	Leu	Gly	Pro	Tyr	Ile	Gln	Met	Ala	Ala	Lys
		1010					1015					1020			
Met	Val	Trp	Ser	Met	Cys	Tyr	Ile	Cys	Val	Leu	Leu	Leu	Val	Pro	Leu
1025					1030					1035					104
Met	Ala	Phe	Gly	Val	Asn	Arg	Gln	Ala	Leu	Thr	Glu	Pro	Asn	Val	Lys
				1045						1050				1055	
Asp	Trp	His	Trp	Leu	Val	Arg	Asn	Ile	Phe	Tyr	Lys	Pro	Tyr	Phe	
			1060				1065					1070			
Met	Leu	Tyr	Gly	Glu	Val	Tyr	Ala	Gly	Glu	Ile	Asp	Thr	Cys	Gly	Asp
		1075					1080					1085			
Glu	Gly	Ile	Arg	Cys	Phe	Pro	Gly	Tyr	Phe	Ile	Pro	Pro	Leu	Leu	Met
		1090				1095					1100				
Val	Ile	Phe	Leu	Leu	Val	Ala	Asn	Ile	Leu	Leu	Leu	Asn	Leu	Leu	Ile
1105					1110					1115					112
Ala	Ile	Phe	Asn	Asn	Ile	Tyr	Asn	Asp	Ser	Ile	Glu	Lys	Ser	Lys	Glu
				1125						1130				1135	
Ile	Trp	Leu	Phe	Gln	Arg	Tyr	Gln	Gln	Leu	Met	Glu	Tyr	His	Asp	Ser
				1140				1145					1150		
Pro	Phe	Leu	Pro	Pro	Pro	Phe	Ser	Ile	Phe	Ala	His	Val	Tyr	His	Phe
		1155					1160					1165			
Ile	Asp	Tyr	Leu	Tyr	Asn	Leu	Arg	Arg	Pro	Asp	Thr	Lys	Arg	Phe	Arg
		1170				1175					1180				
Ser	Glu	His	Ser	Ile	Lys	Leu	Ser	Val	Thr	Glu	Asp	Glu	Met	Lys	Arg
1185					1190					1195					120
Ile	Gln	Asp	Phe	Glu	Glu	Asp	Cys	Ile	Asp	Thr	Leu	Thr	Arg	Ile	Arg
				1205						1210				1215	
Lys	Leu	Lys	Leu	Asn	Thr	Lys	Glu	Pro	Leu	Ser	Val	Thr	Asp	Leu	Thr
			1220				1225					1230			
Glu	Leu	Thr	Cys	Gln	Arg	Val	His	Asp	Leu	Met	Gln	Glu	Asn	Phe	Leu
		1235					1240					1245			
Leu	Lys	Ser	Arg	Val	Tyr	Asp	Ile	Glu	Thr	Lys	Ile	Asp	His	Ile	Ser
		1250				1255					1260				
Asn	Ser	Ser	Asp	Glu	Val	Val	Gln	Ile	Leu	Lys	Asn	Lys	Lys	Leu	Ser
1265				1270						1275					128
Gln	Asn	Phe	Ala	Ala	Ser	Ser	Leu	Ser	Leu	Pro	Asp	Thr	Ser	Ile	Glu
				1285						1290				1295	
Val	Pro	Lys	Ile	Thr	Lys	Thr	Leu	Ile	Asp	Cys	His	Leu	Ser	Pro	Val
			1300					1305				1310			
Ser	Ile	Glu	Asp	Arg	Leu	Ala	Thr	Arg	Ser	Pro	Leu	Leu	Ala	Asn	Leu
		1315					1320					1325			

-25-

Gln Arg Asp His Thr Leu Arg Lys Leu Pro Thr Trp Glu Thr Ser Thr
 1330 1335 1340
 Ala Ser Thr Ser Ser Phe Glu Phe Val Phe Tyr Phe Thr Arg His Glu
 1345 1350 1355 136
 Gly Asn Glu Asn Lys Tyr Glu Phe Lys Lys Leu Glu Lys Gly Gly Phe
 1365 1370 1375
 Trp Arg Asn Asn Tyr Val Ile Ser Trp Arg Leu
 1380 1385

<210> 15
 <211> 1868
 <212> PRT
 <213> C. Elegans

<400> 15
 Met Asn Leu Cys Tyr Arg Arg His Arg Tyr Ala Ser Ser Pro Glu Val
 1 5 10 15
 Trp Cys Thr Met Glu Ser Asp Glu Leu Gly Val Thr Arg Tyr Leu Gln
 20 25 30
 Ser Lys Gly Gly Asp Gln Val Pro Thr Ser Thr Thr Gly Gly
 35 40 45
 Ala Gly Gly Asp Gly Asn Ala Val Pro Thr Thr Ser Gln Ala Gln Ala
 50 55 60
 Gln Thr Phe Asn Ser Gly Arg Gln Thr Thr Gly Met Ser Ser Gly Asp
 65 70 75 80
 Arg Leu Asn Glu Asp Val Ser Ala Thr Ala Asn Ser Ala Gln Leu Val
 85 90 95
 Leu Pro Thr Pro Leu Phe Asn Gln Met Arg Phe Thr Glu Ser Asn Met
 100 105 110
 Ser Leu Asn Arg His Asn Trp Val Arg Glu Thr Phe Thr Arg Arg Glu
 115 120 125
 Cys Ser Arg Phe Ile Ala Ser Ser Arg Asp Leu His Lys Cys Gly Cys
 130 135 140
 Gly Arg Thr Arg Asp Ala His Arg Asn Ile Pro Glu Leu Thr Ser Glu
 145 150 155 160
 Phe Leu Arg Gln Lys Arg Ser Val Ala Ala Leu Glu Gln Gln Arg Ser
 165 170 175
 Ile Ser Asn Val Asn Asp Asp Ile Asn Thr Gln Asn Met Tyr Thr Lys
 180 185 190
 Arg Gly Ala Asn Glu Lys Trp Ser Leu Arg Lys His Thr Val Ser Leu
 195 200 205
 Ala Thr Asn Ala Phe Gly Gln Val Glu Phe Gln Gly Gly Pro His Pro
 210 215 220
 Tyr Lys Ala Gln Tyr Val Arg Val Asn Phe Asp Thr Glu Pro Ala Tyr
 225 230 235 240
 Ile Met Ser Leu Phe Glu His Val Trp Gln Ile Ser Pro Pro Arg Leu
 245 250 255
 Ile Ile Thr Val His Gly Gly Thr Ser Asn Phe Asp Leu Gln Pro Lys
 260 265 270
 Leu Ala Arg Val Phe Arg Lys Gly Leu Leu Lys Ala Ala Ser Thr Thr
 275 280 285
 Gly Ala Trp Ile Ile Thr Ser Gly Cys Asp Thr Gly Val Val Lys His
 290 295 300
 Val Ala Ala Ala Leu Glu Gly Ala Gln Ser Ala Gln Arg Asn Lys Ile
 305 310 315 320
 Val Cys Ile Gly Ile Ala Pro Trp Gly Leu Leu Lys Lys Arg Glu Asp
 325 330 335
 Phe Ile Gly Gln Asp Lys Thr Val Pro Tyr Tyr Pro Ser Ser Ser Lys
 340 345 350
 Gly Arg Phe Thr Gly Leu Asn Asn Arg His Ser Tyr Phe Leu Leu Val

	355					360					365				
Asp	Asn	Gly	Thr	Val	Gly	Arg	Tyr	Gly	Ala	Glu	Val	Ile	Leu	Arg	Lys
	370					375					380				
Arg	Leu	Glu	Met	Tyr	Ile	Ser	Gln	Lys	Gln	Lys	Ile	Phe	Gly	Gly	Thr
385					390					395					400
Arg	Ser	Val	Pro	Val	Val	Cys	Val	Val	Leu	Glu	Gly	Gly	Ser	Cys	Thr
				405					410					415	
Ile	Arg	Ser	Val	Leu	Asp	Tyr	Val	Thr	Asn	Val	Pro	Arg	Val	Pro	Val
			420					425					430		
Val	Val	Cys	Asp	Gly	Ser	Gly	Arg	Ala	Ala	Asp	Leu	Leu	Ala	Phe	Ala
			435				440					445			
His	Gln	Asn	Val	Thr	Glu	Asp	Gly	Leu	Leu	Pro	Asp	Asp	Ile	Arg	Arg
	450					455					460				
Gln	Val	Leu	Leu	Leu	Val	Glu	Thr	Thr	Phe	Gly	Cys	Ser	Glu	Ala	Ala
465					470				475						480
Ala	His	Arg	Leu	Leu	His	Glu	Leu	Thr	Val	Cys	Ala	Gln	His	Lys	Asn
				485					490					495	
Leu	Leu	Thr	Ile	Phe	Arg	Leu	Gly	Glu	Gln	Gly	Glu	His	Asp	Val	Asp
			500					505					510		
His	Ala	Ile	Leu	Thr	Ala	Leu	Leu	Lys	Gly	Gln	Asn	Leu	Ser	Ala	Ala
		515					520					525			
Asp	Gln	Leu	Ala	Leu	Ala	Leu	Ala	Trp	Asn	Arg	Val	Asp	Ile	Ala	Arg
	530					535					540				
Ser	Asp	Val	Phe	Ala	Met	Gly	His	Glu	Trp	Pro	Gln	Ala	Ala	Leu	His
545					550					555					560
Asn	Ala	Met	Met	Glu	Ala	Leu	Ile	His	Asp	Arg	Val	Asp	Phe	Val	Arg
				565					570					575	
Leu	Leu	Leu	Glu	Gln	Gly	Ile	Asn	Met	Gln	Lys	Phe	Leu	Thr	Ile	Ser
			580					585					590		
Arg	Leu	Asp	Glu	Leu	Tyr	Asn	Thr	Asp	Lys	Gly	Pro	Pro	Asn	Thr	Leu
		595					600					605			
Phe	Tyr	Ile	Val	Arg	Asp	Val	Val	Arg	Val	Arg	Gln	Gly	Tyr	Arg	Phe
	610					615					620				
Lys	Leu	Pro	Asp	Ile	Gly	Leu	Val	Ile	Glu	Lys	Leu	Met	Gly	Asn	Ser
625					630					635					640
Tyr	Gln	Cys	Ser	Tyr	Thr	Thr	Ser	Glu	Phe	Arg	Asp	Lys	Tyr	Lys	Gln
				645					650					655	
Arg	Met	Lys	Arg	Val	Lys	His	Ala	Gln	Lys	Lys	Ala	Met	Gly	Val	Phe
			660					665					670		
Ser	Ser	Arg	Pro	Ser	Arg	Thr	Gly	Ser	Gly	Ile	Ala	Ser	Arg	Gln	Ser
		675					680					685			
Thr	Glu	Gly	Met	Gly	Gly	Val	Gly	Gly	Gly	Ser	Ser	Val	Ala	Gly	Val
	690					695						700			
Phe	Gly	Asn	Ser	Phe	Gly	Asn	Gln	Asp	Pro	Pro	Leu	Asp	Pro	His	Val
705					710					715					

-27-

835					840					845				
Ala	Glu	Phe	Arg	Ile	Leu	Ser	Leu	Glu	Leu	Leu	Asp	His	Cys	Tyr
850					855					860				
His	Val	Asp	Asp	Ala	Gln	Thr	Leu	Gln	Leu	Leu	Thr	Tyr	Glu	Leu
865					870					875				
Asn	Trp	Ser	Asn	Glu	Thr	Cys	Leu	Ala	Leu	Ala	Val	Ile	Val	Asn
Lys	His	Phe	Leu	Ala	His	Pro	Cys	Cys	Gln	Ile	Leu	Leu	Ala	Asp
Trp	His	Gly	Gly	Leu	Arg	Met	Arg	Thr	His	Ser	Asn	Ile	Lys	Val
Leu	Gly	Leu	Ile	Cys	Pro	Pro	Phe	Ile	Gln	Met	Leu	Glu	Phe	Lys
Arg	Glu	Glu	Leu	Leu	Asn	Gln	Pro	Gln	Thr	Ala	Ala	Glu	His	Gln
Asp	Met	Asn	Tyr	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser	Ser
Ser	Ser	Ser	Ser	Ser	Asp	Ser	Ser	Ser	Phe	Glu	Asp	Asp	Asp	Glu
Asn	Asn	Ala	His	Asn	His	Asp	Gln	Lys	Arg	Thr	Arg	Lys	Thr	Ser
Gly	Ser	Ala	Gln	Ser	Leu	Asn	Ile	Thr	Ser	Leu	Phe	His	Ser	Arg
Arg	Lys	Ala	Lys	Lys	Asn	Glu	Lys	Cys	Asp	Arg	Glu	Thr	Asp	Ala
Ala	Cys	Glu	Ala	Gly	Asn	Arg	Gln	Ile	Gln	Asn	Gly	Gly	Leu	Thr
Glu	Tyr	Gly	Thr	Phe	Gly	Glu	Ser	Asn	Gly	Val	Ser	Pro	Pro	Pro
Tyr	Met	Arg	Ala	Asn	Ser	Arg	Ser	Arg	Tyr	Asn	Asn	Arg	Ser	Asp
Ser	Lys	Thr	Ser	Ser	Val	Ile	Phe	Gly	Ser	Asp	Pro	Asn	Leu	Ser
Leu	Gln	Lys	Ser	Asn	Ile	Thr	Ser	Thr	Asp	Arg	Pro	Asn	Pro	Met
Gln	Phe	Gln	Gly	Thr	Arg	Lys	Ile	Lys	Met	Arg	Arg	Arg	Phe	Tyr
Phe	Tyr	Ser	Ala	Pro	Ile	Ser	Thr	Phe	Trp	Ser	Trp	Thr	Ile	Ser
Ile	Leu	Phe	Ile	Thr	Phe	Phe	Thr	Tyr	Thr	Leu	Leu	Val	Lys	Thr
Pro	Arg	Pro	Thr	Val	Ile	Glu	Tyr	Ile	Leu	Ile	Ala	Tyr	Val	Ala
Phe	Gly	Leu	Glu	Gln	Val	Arg	Lys	Ile	Ile	Met	Ser	Asp	Ala	Lys
Phe	Tyr	Glu	Lys	Ile	Arg	Thr	Tyr	Val	Cys	Ser	Phe	Trp	Asn	Cys
Thr	Ile	Leu	Ala	Ile	Ile	Phe	Tyr	Ile	Val	Gly	Phe	Phe	Met	Arg
Phe	Gly	Ser	Val	Ala	Tyr	Gly	Arg	Val	Ile	Leu	Ala	Cys	Asp	Ser
Leu	Trp	Thr	Met	Lys	Leu	Leu	Asp	Tyr	Met	Ser	Val	His	Pro	Lys
Gly	Pro	Tyr	Val	Thr	Met	Ala	Gly	Lys	Met	Ile	Gln	Asn	Met	Ser
Ile	Ile	Val	Met	Leu	Val	Val	Thr	Leu	Leu	Ser	Phe	Gly	Leu	Ala
Gln	Ser	Ile	Thr	Tyr	Pro	Asp	Glu	Thr	Trp	His	Trp	Ile	Leu	Val
Asn	Ile	Phe	Leu	Lys	Pro	Tyr	Phe	Met	Leu	Tyr	Gly	Glu	Val	Tyr

1315	1320	1325
Asp Glu Ile Asp Thr Cys Gly Asp Glu Ala Trp Asp Gln His Leu Glu		
1330	1335	1340
Asn Gly Gly Pro Val Ile Leu Gly Asn Gly Thr Thr Gly Leu Ser Cys		
1345	1350	1355
Val Pro Gly Tyr Trp Ile Pro Pro Leu Leu Met Thr Phe Phe Leu Leu		
1365	1370	1375
Ile Ala Asn Ile Leu Leu Met Ser Met Leu Ile Ala Ile Phe Asn His		
1380	1385	1390
Ile Phe Asp Ala Thr Asp Glu Met Ser Gln Gln Ile Trp Leu Phe Gln		
1395	1400	1405
Arg Tyr Lys Gln Val Met Glu Tyr Glu Ser Thr Pro Phe Leu Pro Pro		
1410	1415	1420
Pro Leu Thr Pro Leu Tyr His Gly Val Leu Ile Leu Gln Phe Val Arg		
1425	1430	1435
Thr Arg Leu Ser Cys Ser Lys Ser Gln Glu Arg Asn Pro Ile Leu Leu		
1445	1450	1455
Leu Lys Ile Ala Glu Leu Phe Leu Asp Asn Asp Gln Ile Glu Lys Leu		
1460	1465	1470
His Asp Phe Glu Glu Asp Cys Met Glu Asp Leu Ala Arg Gln Lys Leu		
1475	1480	1485
Asn Glu Lys Asn Thr Ser Asn Glu Gln Arg Ile Leu Arg Ala Asp Ile		
1490	1495	1500
Arg Thr Asp Gln Ile Leu Asn Arg Leu Ile Asp Leu Gln Ala Lys Glu		
1505	1510	1515
Ser Met Gly Arg Asp Val Ile Asn Asp Val Glu Ser Arg Leu Ala Ser		
1525	1530	1535
Val Glu Lys Ala Gln Asn Glu Ile Leu Glu Cys Val Arg Ala Leu Leu		
1540	1545	1550
Asn Gln Asn Asn Ala Pro Thr Ala Ile Gly Arg Cys Phe Ser Pro Ser		
1555	1560	1565
Pro Asp Pro Leu Val Glu Thr Ala Asn Gly Thr Pro Gly Pro Leu Leu		
1570	1575	1580
Leu Lys Leu Pro Gly Thr Asp Pro Ile Leu Glu Glu Lys Asp His Asp		
1585	1590	1595
Ser Gly Glu Asn Ser Asn Ser Leu Pro Pro Gly Arg Ile Arg Arg Asn		
1605	1610	1615
Arg Thr Ala Thr Ile Cys Gly Gly Tyr Val Ser Glu Glu Arg Asn Met		
1620	1625	1630
Met Leu Leu Ser Pro Lys Pro Ser Asp Val Ser Gly Ile Pro Gln Gln		
1635	1640	1645
Arg Leu Met Ser Val Thr Ser Met Asp Pro Leu Pro Leu Pro Leu Ala		
1650	1655	1660
Lys Leu Ser Thr Met Ser Ile Arg Arg Arg His Glu Glu Tyr Thr Ser		
1665	1670	1675
Ile Thr Asp Ser Ile Ala Ile Arg His Pro Glu Arg Arg Ile Arg Asn		
1685	1690	1695
Asn Arg Ser Asn Ser Ser Glu His Asp Glu Ser Ala Val Asp Ser Glu		
1700	1705	1710
Gly Gly Gly Asn Val Thr Ser Ser Pro Arg Lys Arg Ser Thr Arg Asp		
1715	1720	1725
Leu Arg Met Thr Pro Ser Ser Gln Val Glu Glu Ser Thr Ser Arg Asp		
1730	1735	1740
Gln Ile Phe Glu Ile Asp His Pro Glu His Glu Glu Asp Glu Ala Gln		
1745	1750	1755
Ala Asp Cys Glu Leu Thr Asp Val Ile Thr Glu Glu Glu Asp Glu Glu		
1765	1770	1775
Glu Asp Asp Glu Glu Asp Asp Ser His Glu Arg His His Ile His Pro		
1780	1785	1790
Arg Arg Lys Ser Ser Arg Gln Asn Arg Gln Pro Ser His Thr Leu Glu		

-29-

1795 1800 1805
 Thr Asp Leu Ser Glu Gly Glu Glu Val Asp Pro Leu Asp Val Leu Lys
 1810 1815 1820
 Met Lys Glu Leu Pro Ile Ile His Gln Ile Leu Asn Glu Glu Glu Gln
 1825 1830 1835 184
 Ala Gly Ala Pro His Ser Thr Pro Val Ile Ala Ser Pro Ser Ser Ser
 1845 1850 1855
 Arg Ala Asp Leu Thr Ser Gln Lys Cys Ser Asp Val
 1860 1865

<210> 16
 <211> 489
 <212> DNA
 <213> Mus Musculus

<400> 16
 ccctgaaaga ctcgacttct gctgctagcg ctggagctga gttagttttg agaaggtttc 60
 ccggggctgt ccttggttcgg tggcccgtgc caccgcctcc ggagacgctt tccgatagat 120
 ggctgcaggc cgcggagggtg gaggaggagc cgtgcccctt ccggagtccg ccccgtagag 180
 agaattgtccc agaaatcctg gatagagagc actttgacca agagggagtg tgtatatatt 240
 ataccaagct ccaaagaccc tcacagatgt cttccaggat gtcagatttg tcagcaactt 300
 gtcagatggt tctgtggtcg tttggtcaag caacatgcat gctttactgc aagtcttgcc 360
 atgaaatact cagatgtgaa attgggtgaa cactttaacc aggcaataga agaatggtct 420
 gtgaaaagc acacggagca gagcccaaca gatgcttatg gagtcatcaa ttttcaaggg 480
 ggttctcat 489

<210> 17
 <211> 102
 <212> PRT
 <213> Mus Musculus

<400> 17
 Met Ser Gln Lys Ser Trp Ile Glu Ser Thr Leu Thr Lys Arg Glu Cys
 1 5 10 15
 Val Tyr Ile Ile Pro Ser Ser Lys Asp Pro His Arg Cys Leu Pro Gly
 20 25 30
 Cys Gln Ile Cys Gln Gln Leu Val Arg Cys Phe Cys Gly Arg Leu Val
 35 40 45
 Lys Gln His Ala Cys Phe Thr Ala Ser Leu Ala Met Lys Tyr Ser Asp
 50 55 60
 Val Lys Leu Gly Glu His Phe Asn Gln Ala Ile Glu Glu Trp Ser Val
 65 70 75 80
 Glu Lys His Thr Glu Gln Ser Pro Thr Asp Ala Tyr Gly Val Ile Asn
 85 90 95
 Phe Gln Gly Gly Ser His
 100

<210> 18
 <211> 410
 <212> DNA
 <213> Homo Sapiens

<220>
 <221> unsure
 <222> (6)...(6)
 <221> unsure
 <222> (58)...(58)
 <221> unsure

-30-

<222> (89)...(89)

<221> unsure

<222> (406)...(406)

<400> 18

gccgcnggag	cctgagcgga	gggtgtgcgc	agcctcgcca	gcggggggccc	cgggctgngc	60
cattgcctca	ctgagccagc	gcctgcctnc	tacctcgccg	acagctggaa	ccagtgcgac	120
ctagtggctc	tcacctgctt	cctcctgggc	gtgggctgcc	ggctgacccc	gggtttgtac	180
cacctggggc	gcactgtcct	ctgcatcgac	ttcatggttt	tcacggtgcg	gctgcttcac	240
atcttcacgg	tcaacaaaca	gctggggccc	aagatcgta	tcgtgagcaa	gatgatgaag	300
gacgtgttct	tcttcctctt	cttcctcggc	gtgtggctgg	tagctatggg	ttgggccacg	360
gaggggttcc	tgaggccacg	ggacagtgc	ttcccaagta	tcctgncgcc		410

<210> 19

<211> 131

<212> PRT

<213> Homo Sapiens

<220>

<221> UNSURE

<222> (15)...(15)

<223> UNKNOWN

<221> UNSURE

<222> (25)...(25)

<223> UNKNOWN

<221> UNSURE

<222> (131)...(131)

<223> UNKNOWN

<400> 19

Ala	Glu	Gly	Val	Arg	Ser	Leu	Ala	Ser	Gly	Gly	Pro	Gly	Leu	Xaa	His
1				5					10					15	
Cys	Leu	Thr	Glu	Pro	Ala	Pro	Ala	Xaa	Tyr	Leu	Ala	Asp	Ser	Trp	Asn
			20					25					30		
Gln	Cys	Asp	Leu	Val	Ala	Leu	Thr	Cys	Phe	Leu	Leu	Gly	Val	Gly	Cys
		35					40					45			
Arg	Leu	Thr	Pro	Gly	Leu	Tyr	His	Leu	Gly	Arg	Thr	Val	Leu	Cys	Ile
	50					55					60				
Asp	Phe	Met	Val	Phe	Thr	Val	Arg	Leu	Leu	His	Ile	Phe	Thr	Val	Asn
	65				70					75					80
Lys	Gln	Leu	Gly	Pro	Lys	Ile	Val	Ile	Val	Ser	Lys	Met	Met	Lys	Asp
			85						90					95	
Val	Phe	Phe	Phe	Leu	Phe	Phe	Leu	Gly	Val	Trp	Leu	Val	Ala	Met	Gly
			100					105					110		
Trp	Ala	Thr	Glu	Gly	Phe	Leu	Arg	Pro	Arg	Asp	Ser	Asp	Phe	Pro	Ser
		115					120					125			
Ile	Leu	Xaa													
			130												

<210> 20

<211> 389

<212> DNA

<213> Homo Sapiens

<400> 20

caaatttttt	gttagtacac	catctcatcc	aaattgcaaa	agtcacatgg	aaactggaac	60
caaagatcaa	gaaactgttt	gctctaaagc	tacagaagga	gataatacag	aatttgagac	120

-31-

at ttg tagga	cac agagata	gcat ggatt	ac agagg ttt	aa agaaacat	caaacaagat	180
aaaaatacta	tccaataaca	atacttctga	aaacactttg	aaacgagtga	gttctcttgc	240
tggatttact	gactgtcaca	gaacttccat	tcctgttcat	tcaaaacgag	aaaagatcag	300
tagaaggcca	tctaccgaag	acactcatga	agtagattcc	aaagcagctt	taataccggt	360
ttgtagattt	caactaaaca	gatatatat				389

<210> 21
 <211> 415
 <212> DNA
 <213> Homo Sapiens

<400> 21						
at ttctagtt	tttcaaattt	gccagtcttt	ttgaatagta	tctccttctt	ttctcatggt	60
ttatat ttaa	aactttttta	tgtccatcat	cacttttaa	at actttattt	tgtcatctat	120
aaccaataat	tccactatct	tatcagaaat	caaataccgt	ttatgtaagt	tgactcccat	180
gagttctaaa	ttgccattgt	gaggtcatct	tcggttaggc	tttaatttgt	tgcaaagttg	240
tgca gctcag	ggtcagggaag	agtccttcca	gaaaggagga	tttgttactg	tgaatctctt	300
tg ttaactaa	cctctttccc	cactgaaata	acttttttca	ataacatgat	tttaacaaca	360
taatctctct	atgccagaac	agatatatat	gaatgtaagt	caatattttc	ttgag	415

<210> 22
 <211> 405
 <212> DNA
 <213> Mus Musculus

<400> 22						
ttattatggc	ttatcatgaa	aaaccagtcc	tgcctcctcc	tcttatcatc	ctcagccata	60
tagttt cact	gttttgctgt	gtatgcaaaa	gaagaaagaa	agataagact	tccgatgggc	120
caaaactttt	cttaacagaa	gaagatcaaa	agaaactcca	tgattttgaa	gagcagtgtg	180
ttgagatgta	ctttgatgag	aaagatgaca	aattcaattc	tgggagtgaa	gagagaatcc	240
gggtcacttt	tgaaagagtg	gagcagatga	gcattcagat	taaagaagtt	ggagatcgtg	300
tcaactacat	aaaaagatca	ttacagtctt	tagattctca	aattggtcat	ctgcaagatc	360
tctcagccct	aacagtagat	acattgaaaa	cacttacagc	ccaga		405

<210> 23
 <211> 5117
 <212> DNA
 <213> Homo Sapiens

<220>
 <221> unsure
 <222> (2382) ... (2382)
 <223> unknown

<221> unsure
 <222> (4664) ... (4664)
 <223> unknown

<221> unsure
 <222> (4682) ... (4682)
 <223> unknown

<221> unsure
 <222> (4702) ... (4702)
 <223> unknown

<221> unsure
 <222> (5038) ... (5039)
 <223> unknown

<221> unsure
 <222> (5056) ... (5056)
 <223> unknown

<221> unsure
 <222> (5071) ... (5072)

<400> 23

gatggcaaca	tgggtgaagaa	tcaatggcta	aagcattagt	tgccctgtaag	atctatcggt	60
caatggcata	tgaagcaaa	cagagtgaac	tggtagatga	tacttcagaa	gaactaaaac	120
agtattccaa	tgattttggt	cagttggccg	ttgaattatt	agaacagtcc	ttcagacaag	180
atgaaaccaa	ggctatgaaa	ttgctcactt	atgaactgaa	gaactggagt	aattcaacct	240
gccttaagtt	agcagtttct	tcaagactta	gaccttttgt	agctcacacc	tgtacacaaa	300
tgttgttatc	tgatatgtgg	atgggaaggc	tgaatatgag	gaaaaattcc	tggtacaagg	360
tcatactaag	catttttagtt	ccacctgcca	tattgctggt	agagtataaa	actaaggctg	420
aaatgtccca	tatcccacaa	tctcaagatg	ctcatcagat	gacaatggat	gacagcgaaa	480
acaactttca	gaacataaca	gaagagatcc	ccatggaagt	gtttaaagaa	gtacggattt	540
tgatagtaa	tgaaggaaa	aatgagatgg	agatacaaat	gaaatcaaaa	aagcttccaa	600
ttacgcgaaa	gttttatgcc	ttttatcatg	caccaattgt	aaaattctgg	tttaacacgt	660
tggtcatatt	aggatttctg	atgctttata	catttgtggt	tcttgtacaa	atggaacagt	720
taccttcagt	tcaagaatgg	attgttattg	cttatatttt	tacttatgcc	attgagaaa	780
tcgtgagat	ctttatgtct	gaagctggga	aagtaaacca	gaagattaaa	gtatggttta	840
gtgattactt	caacatcagt	gatacaattg	ccataatttc	tttcttcatt	ggatttggac	900
taagatttgg	agcaaaatgg	aactttgcaa	atgcatatga	taatcatggt	tttgtggctg	960
gaagattaat	ttactgtctt	aacataatat	tttggtatgt	gcgtttgcta	gattttctag	1020
ctgtaaatca	acaggcgagg	ccttatgtaa	tgatgattgg	aaaaatggtg	gccaatatgt	1080
tctacattgt	agtgtattat	gctcttgat	tacttagttt	tggtgttccc	agaaaggcaa	1140
tactttatcc	tcatgaagca	ccatcttgga	ctcttgctaa	agatatagtt	tttcaacct	1200
actggatgat	ttttggtgaa	gtttatgcat	acgaaattga	tgtgtgtgca	aatgattctg	1260
ttatccctca	aatctgtggt	cctgggacgt	ggttgactcc	atttcttcaa	gcagtctacc	1320
tctttgtaca	gtatatcatt	atgggttaatc	ttcttattgc	atttttcaac	aatgtgtatt	1380
tacaagtga	ggcaatttcc	aatattgtat	ggaagtacca	gcgttatcat	tttattatgg	1440
cttatcatga	gaaaccagtt	ctgcctcctc	cacttatcat	tcttagccat	atagtttctc	1500
tgttttgctg	catatgtaag	agaagaaaga	aagataagac	ttccgatgga	ccaaaacttt	1560
tcttaacaga	agaagatcaa	aagaaacttc	atgattttga	agagcagttg	gttgaaatgt	1620
atttcaatga	aaaagatgac	aaatttcatt	ctgggagtga	agagagaatt	cgtgtcactt	1680
ttgaaagagt	ggaacagatg	tgcattcaga	ttaaagaagt	tgagatcggt	gtcaactaca	1740
taaaaaagatc	attacaatca	ttagattctc	aaattggcca	tttgcaagat	ctttcagccc	1800
tgacggtaga	tacattaaaa	acactcactg	cccagaaagc	gtcggaagct	agcaaagttc	1860
ataatgaaat	cacacgagaa	ctgagcattt	ccaaacactt	ggctcaaaac	cttattgatg	1920
atggtcctgt	aagaccttct	gtatggaaaa	agcatgggtg	tgtaaataca	cttagctcct	1980
ctcttcctca	aggtgatctt	gaaagtaata	atccttttca	ttgtaatat	ttaatgaaag	2040
atgacaaaag	tccccagttg	aatatatatt	gtcaagactt	acctgcagta	ccccagagaa	2100
aagaatttaa	ttttccagag	gctggttcct	cttctggtgc	cttattccca	agtgtgtgtt	2160
cccctccaga	actgcgacag	agactacatg	gggtagaact	cttaaaaaata	tttaataaaa	2220
atcaaaaatt	aggcagttca	tctactagca	taccacatct	gtcatcccca	ccaaccaa	2280
ttttgttag	tacaccatct	cagccaagtt	gcaaaagcca	cttggaact	ggaaccaa	2340
atcaagaaac	tgtttgcctt	aaagctacag	aaggagataa	tncagaattt	ggagcatttg	2400
taggacacag	agatagcatg	gatttacaga	ggtttaaaga	aacatcaaac	aagataaaaa	2460
tactatccaa	taacaatact	tctgaaaaca	ctttgaaacg	agtgagttct	cttgcctggat	2520
ttactgactg	tcacagaact	tccattcctg	ttcattcaaa	acaagcagaa	aaaatcagta	2580
gaaggccatc	taccgaagac	actcatgaag	tagattccaa	agcagcttta	ataccggatt	2640
ggttacaaga	tagaccatca	aacagagaaa	tgccatctga	agaaggaaca	ttaaatggtc	2700
tcacttctcc	atttaagcca	gctatggata	caaattacta	ttattcagct	gtggaaagaa	2760
ataacttgat	gaggttatca	cagagcattc	catttacacc	tgtgcctcca	agaggggagc	2820
ctgtcacagt	gtatcgtttg	gaagagagtt	cacccaacat	actaaataac	agcatgtctt	2880
cttggtcaca	actaggcctc	tgtgccaaaa	tagagttttt	aagcaaagag	gagatgggag	2940
gaggtttacg	aagagctgtc	aaagtacagt	gtacctggtc	agaacatgat	atcctcaaat	3000
cagggcatct	ttatattatc	aaatcttttc	ttccagaggt	ggttaataca	tggtcaagta	3060
tttataaaga	agatacagtt	ctgcatctct	gtctgagaga	aattcaacaa	cagagagcag	3120

-33-

```

cacaaaagct tacgtttgcc tttaatcaaa tgaaacccaa atccatacca tattctccaa 3180
ggttccttga agttttcctg ctgtattgcc attcagcagg acagtgggtt gctgtggaag 3240
aatgtatgac tggagaattt agaaaataca acaataataa tggagatgag attattccaa 3300
ctaatactct ggaagagatc atgctagcct ttagccactg gacttacgaa tatacaagag 3360
gggagtactt ggtacttgat ttgcaagggt ttggtgaaaa ttgactgac ccatctgtga 3420
taaaagcaga agaaaagaga tcctgtgata tgggttttgg cccagcaaat ctaggagaag 3480
atgcaattaa aaacttcaga gcaaaacatc actgtaattc ttgctgtaga aagcttaaac 3540
ttccagatct gaagaggaat gattatacgc ctgataaaat tatatttcct caggatgagc 3600
cttcagattt gaatcttcag cctggaaatt ccaccaaaaga atcagaatca gctaattctg 3660
ttcgtctgat gttataatat taatattact gaatcattgg ttttgcctgc acctcacaga 3720
aatgttactg tgtcactttt ccctcgggag gaaattgttt ggtaatatag aaaggtgtat 3780
gcaagttgaa tttgctgact ccagcacagt taaaagggtc atattctttt gacctgatta 3840
atcagtcaga aagtccttat aggatagagc tggcagctga gaaattttaa aggtaattga 3900
taattagtag ttgtaacttt ttaaagggtt ctttgtatag cagaggatct catttgactt 3960
tgttttgatg aggggtgatgc cctctcttat gtggtacaat accattaacc aaaggtaggt 4020
gtccatgcag attttatttg cagctgtttt attgccattc aactagggaa atgaagaaat 4080
cacgcagcct tttggtttaa tggcagctca aattttcctc agtgtattta gtgtgttcag 4140
tgatgatata actgtttccc aactagatgc ttgttggtcca cgggaaggga aatgacttgt 4200
tctaattcta ggttcacaga ggtatgagaa gcctgaactg aagaccattt tcaagaggga 4260
cggatattat gaatcagggt taggctccat atttaaagat agagccagtt ttttttttaa 4320
atagaaccca aattgtgtaa aaatgttaat tgggtttttt aaacattgtt ttatcaagtc 4380
actgttaagt agaagaaaagc catggtaaac tgatacataa cctaaattat aaaagcagaa 4440
acctaaactc ctctgcaagg gaagttacct tttgaggaaa gttaaagtac ttttttccct 4500
atctgtatct atagcaacaa cccagaactt acaaacttct ccaaagattt tattgattgt 4560
tatatcaaat cagaatgtaa acatgaactc ttgcatatat ttaaaattgt gttggaacat 4620
ttgaacatga atgctgtttg ggtacttaag aaattrattc agtnggatta tcattatgtg 4680
anactggcag attgcagtgc ancttatgc caataaaatg taatttaaca gccccagata 4740
ttgttgaata ttcaacaata acaagaaaag cttttcatct aagttttatg ctttaatttt 4800
ttttcttttt ttttcttttt cttttgtttc cttggtacta attttaattt ttatttgtaa 4860
gggagcagta taaagcttat ttgtatttag tagtgtatct catagatata gacaaggcaa 4920
gagatgataa cctgttttaa tagtgtttaa tattgattgg ggggtggggag aaagaaaaag 4980
tgtattactt aaagatacta tatacgtttt gtatatcatt aaatctttta aagaaatnna 5040
ataaatttat tgtttncaaa aaaaaaaccc nntaaaaaaa aaagggcggc ccctctagag 5100
gatccctcga ggggcccc 5117

```

<210> 24
 <211> 1224
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> UNSURE
 <222> (794) ... (794)
 <223> UNKNOWN

<400> 24
 Trp Gln His Gly Glu Glu Ser Met Ala Lys Ala Leu Val Ala Cys Lys
 1 5 10 15
 Ile Tyr Arg Ser Met Ala Tyr Glu Ala Lys Gln Ser Asp Leu Val Asp
 20 25 30
 Asp Thr Ser Glu Glu Leu Lys Gln Tyr Ser Asn Asp Phe Gly Gln Leu
 35 40 45
 Ala Val Glu Leu Leu Glu Gln Ser Phe Arg Gln Asp Glu Thr Met Ala
 50 55 60
 Met Lys Leu Leu Thr Tyr Glu Leu Lys Asn Trp Ser Asn Ser Thr Cys
 65 70 75 80
 Leu Lys Leu Ala Val Ser Ser Arg Leu Arg Pro Phe Val Ala His Thr
 85 90 95

-34-

Cys	Thr	Gln	Met	Leu	Leu	Ser	Asp	Met	Trp	Met	Gly	Arg	Leu	Asn	Met
			100					105					110		
Arg	Lys	Asn	Ser	Trp	Tyr	Lys	Val	Ile	Leu	Ser	Ile	Leu	Val	Pro	Pro
		115					120					125			
Ala	Ile	Leu	Leu	Leu	Glu	Tyr	Lys	Thr	Lys	Ala	Glu	Met	Ser	His	Ile
	130					135					140				
Pro	Gln	Ser	Gln	Asp	Ala	His	Gln	Met	Thr	Met	Asp	Asp	Ser	Glu	Asn
145				150						155				160	
Asn	Phe	Gln	Asn	Ile	Thr	Glu	Glu	Ile	Pro	Met	Glu	Val	Phe	Lys	Glu
			165						170					175	
Val	Arg	Ile	Leu	Asp	Ser	Asn	Glu	Gly	Lys	Asn	Glu	Met	Glu	Ile	Gln
		180						185					190		
Met	Lys	Ser	Lys	Lys	Leu	Pro	Ile	Thr	Arg	Lys	Phe	Tyr	Ala	Phe	Tyr
		195				200						205			
His	Ala	Pro	Ile	Val	Lys	Phe	Trp	Phe	Asn	Thr	Leu	Ala	Tyr	Leu	Gly
	210				215						220				
Phe	Leu	Met	Leu	Tyr	Thr	Phe	Val	Val	Leu	Val	Gln	Met	Glu	Gln	Leu
225				230						235				240	
Pro	Ser	Val	Gln	Glu	Trp	Ile	Val	Ile	Ala	Tyr	Ile	Phe	Thr	Tyr	Ala
			245						250					255	
Ile	Glu	Lys	Val	Arg	Glu	Ile	Phe	Met	Ser	Glu	Ala	Gly	Lys	Val	Asn
		260						265						270	
Gln	Lys	Ile	Lys	Val	Trp	Phe	Ser	Asp	Tyr	Phe	Asn	Ile	Ser	Asp	Thr
		275				280						285			
Ile	Ala	Ile	Ile	Ser	Phe	Phe	Ile	Gly	Phe	Gly	Leu	Arg	Phe	Gly	Ala
	290					295					300				
Lys	Trp	Asn	Phe	Ala	Asn	Ala	Tyr	Asp	Asn	His	Val	Phe	Val	Ala	Gly
305				310					315					320	
Arg	Leu	Ile	Tyr	Cys	Leu	Asn	Ile	Ile	Phe	Trp	Tyr	Val	Arg	Leu	Leu
			325						330					335	
Asp	Phe	Leu	Ala	Val	Asn	Gln	Gln	Ala	Gly	Pro	Tyr	Val	Met	Met	Ile
		340						345					350		
Gly	Lys	Met	Val	Ala	Asn	Met	Phe	Tyr	Ile	Val	Val	Ile	Met	Ala	Leu
		355					360					365			
Val	Leu	Leu	Ser	Phe	Gly	Val	Pro	Arg	Lys	Ala	Ile	Leu	Tyr	Pro	His
	370					375					380				
Glu	Ala	Pro	Ser	Trp	Thr	Leu	Ala	Lys	Asp	Ile	Val	Phe	His	Pro	Tyr
385				390					395					400	
Trp	Met	Ile	Phe	Gly	Glu	Val	Tyr	Ala	Tyr	Glu	Ile	Asp	Val	Cys	Ala
			405						410					415	
Asn	Asp	Ser	Val	Ile	Pro	Gln	Ile	Cys	Gly	Pro	Gly	Thr	Trp	Leu	Thr
		420						425					430		
Pro	Phe	Leu	Gln	Ala	Val	Tyr	Leu	Phe	Val	Gln	Tyr	Ile	Ile	Met	Val
		435				440						445			
Asn	Leu	Leu	Ile	Ala	Phe	Phe	Asn	Asn	Val	Tyr	Leu	Gln	Val	Lys	Ala
	450					455					460				
Ile	Ser	Asn	Ile	Val	Trp	Lys	Tyr	Gln	Arg	Tyr	His	Phe	Ile	Met	Ala
465				470					475					480	
Tyr	His	Glu	Lys	Pro	Val	Leu	Pro	Pro	Pro	Leu	Ile	Ile	Leu	Ser	His
			485						490					495	
Ile	Val	Ser	Leu	Phe	Cys	Cys	Ile	Cys	Lys	Arg	Arg	Lys	Lys	Asp	Lys
		500						505					510		
Thr	Ser	Asp	Gly	Pro	Lys	Leu	Phe	Leu	Thr	Glu	Glu	Asp	Gln	Lys	Lys
		515				520						525			
Leu	His	Asp	Phe	Glu	Glu	Gln	Cys	Val	Glu	Met	Tyr	Phe	Asn	Glu	Lys
	530					535					540				
Asp	Asp	Lys	Phe	His	Ser	Gly	Ser	Glu	Glu	Arg	Ile	Arg	Val	Thr	Phe
545				550					555					560	
Glu	Arg	Val	Glu	Gln	Met	Cys	Ile	Gln	Ile	Lys	Glu	Val	Gly	Asp	Arg
			565					570						575	

-35-

Val	Asn	Tyr	Ile	Lys	Arg	Ser	Leu	Gln	Ser	Leu	Asp	Ser	Gln	Ile	Gly
			580					585					590		
His	Leu	Gln	Asp	Leu	Ser	Ala	Leu	Thr	Val	Asp	Thr	Leu	Lys	Thr	Leu
		595					600					605			
Thr	Ala	Gln	Lys	Ala	Ser	Glu	Ala	Ser	Lys	Val	His	Asn	Glu	Ile	Thr
	610					615					620				
Arg	Glu	Leu	Ser	Ile	Ser	Lys	His	Leu	Ala	Gln	Asn	Leu	Ile	Asp	Asp
625					630					635					640
Gly	Pro	Val	Arg	Pro	Ser	Val	Trp	Lys	Lys	His	Gly	Val	Val	Asn	Thr
				645						650				655	
Leu	Ser	Ser	Ser	Leu	Pro	Gln	Gly	Asp	Leu	Glu	Ser	Asn	Asn	Pro	Phe
			660					665					670		
His	Cys	Asn	Ile	Leu	Met	Lys	Asp	Asp	Lys	Asp	Pro	Gln	Cys	Asn	Ile
		675					680					685			
Phe	Gly	Gln	Asp	Leu	Pro	Ala	Val	Pro	Gln	Arg	Lys	Glu	Phe	Asn	Phe
	690					695					700				
Pro	Glu	Ala	Gly	Ser	Ser	Ser	Gly	Ala	Leu	Phe	Pro	Ser	Ala	Val	Ser
705					710					715					720
Pro	Pro	Glu	Leu	Arg	Gln	Arg	Leu	His	Gly	Val	Glu	Leu	Leu	Lys	Ile
				725					730					735	
Phe	Asn	Lys	Asn	Gln	Lys	Leu	Gly	Ser	Ser	Ser	Thr	Ser	Ile	Pro	His
			740					745					750		
Leu	Ser	Ser	Pro	Pro	Thr	Lys	Phe	Phe	Val	Ser	Thr	Pro	Ser	Gln	Pro
		755					760					765			
Ser	Cys	Lys	Ser	His	Leu	Glu	Thr	Gly	Thr	Lys	Asp	Gln	Glu	Thr	Val
	770					775					780				
Cys	Ser	Lys	Ala	Thr	Glu	Gly	Asp	Asn	Xaa	Glu	Phe	Gly	Ala	Phe	Val
785					790					795					800
Gly	His	Arg	Asp	Ser	Met	Asp	Leu	Gln	Arg	Phe	Lys	Glu	Thr	Ser	Asn
				805					810					815	
Lys	Ile	Lys	Ile	Leu	Ser	Asn	Asn	Asn	Thr	Ser	Glu	Asn	Thr	Leu	Lys
			820					825					830		
Arg	Val	Ser	Ser	Leu	Ala	Gly	Phe	Thr	Asp	Cys	His	Arg	Thr	Ser	Ile
		835					840					845			
Pro	Val	His	Ser	Lys	Gln	Ala	Glu	Lys	Ile	Ser	Arg	Arg	Pro	Ser	Thr
	850					855					860				
Glu	Asp	Thr	His	Glu	Val	Asp	Ser	Lys	Ala	Ala	Leu	Ile	Pro	Asp	Trp
865					870					875					880
Leu	Gln	Asp	Arg	Pro	Ser	Asn	Arg	Glu	Met	Pro	Ser	Glu	Glu	Gly	Thr
				885					890					895	
Leu	Asn	Gly	Leu	Thr	Ser	Pro	Phe	Lys	Pro	Ala	Met	Asp	Thr	Asn	Tyr
		900						905					910		
Tyr	Tyr	Ser	Ala	Val	Glu	Arg	Asn	Asn	Leu	Met	Arg	Leu	Ser	Gln	Ser
		915					920					925			
Ile	Pro	Phe	Thr	Pro	Val	Pro	Pro	Arg	Gly	Glu	Pro	Val	Thr	Val	Tyr
	930					935					940				
Arg	Leu	Glu	Glu	Ser	Ser	Pro	Asn	Ile	Leu	Asn	Asn	Ser	Met	Ser	Ser
945					950					955					960
Trp	Ser	Gln	Leu	Gly	Leu	Cys	Ala	Lys	Ile	Glu	Phe	Leu	Ser	Lys	Glu
				965					970					975	
Glu	Met	Gly	Gly	Gly	Leu	Arg	Arg	Ala	Val	Lys	Val	Gln	Cys	Thr	Trp
			980					985					990		
Ser	Glu	His	Asp	Ile	Leu	Lys	Ser	Gly	His	Leu	Tyr	Ile	Ile	Lys	Ser
		995					1000					1005			
Phe	Leu	Pro	Glu	Val	Val	Asn	Thr	Trp	Ser	Ser	Ile	Tyr	Lys	Glu	Asp
	1010					1015					1020				
Thr	Val	Leu	His	Leu	Cys	Leu	Arg	Glu	Ile	Gln	Gln	Gln	Arg	Ala	Ala
1025					1030					1035					104
Gln	Lys	Leu	Thr	Phe	Ala	Phe	Asn	Gln	Met	Lys	Pro	Lys	Ser	Ile	Pro
				1045					1050					1055	

-36-

Tyr Ser Pro Arg Phe Leu Glu Val Phe Leu Leu Tyr Cys His Ser Ala
 1060 1065 1070
 Gly Gln Trp Phe Ala Val Glu Glu Cys Met Thr Gly Glu Phe Arg Lys
 1075 1080 1085
 Tyr Asn Asn Asn Asn Gly Asp Glu Ile Ile Pro Thr Asn Thr Leu Glu
 1090 1095 1100
 Glu Ile Met Leu Ala Phe Ser His Trp Thr Tyr Glu Tyr Thr Arg Gly
 1105 1110 1115 112
 Glu Leu Leu Val Leu Asp Leu Gln Gly Val Gly Glu Asn Leu Thr Asp
 1125 1130 1135
 Pro Ser Val Ile Lys Ala Glu Glu Lys Arg Ser Cys Asp Met Val Phe
 1140 1145 1150
 Gly Pro Ala Asn Leu Gly Glu Asp Ala Ile Lys Asn Phe Arg Ala Lys
 1155 1160 1165
 His His Cys Asn Ser Cys Cys Arg Lys Leu Lys Leu Pro Asp Leu Lys
 1170 1175 1180
 Arg Asn Asp Tyr Thr Pro Asp Lys Ile Ile Phe Pro Gln Asp Glu Pro
 1185 1190 1195 120
 Ser Asp Leu Asn Leu Gln Pro Gly Asn Ser Thr Lys Glu Ser Glu Ser
 1205 1210 1215
 Ala Asn Ser Val Arg Leu Met Leu
 1220

<210> 25
 <211> 2180
 <212> DNA
 <213> Homo Sapiens

<400> 25
 tcgaggccaa gaattcggca cgagggcctc gggcaggccc cctggagcga cctgcttctt 60
 tgggcactgt tgctgaacag ggcacagatg gccatgtact tctgggagat gggttccaat 120
 gcagtttctt cagctcttgg ggcctgtttg ctgctccggg tgatggcagc cctggagcct 180
 gacgctgagg aggcagcacg gaggaaagac ctggcggttca agtttgaggg gatgggcgtt 240
 gacctctttg gcgagtgtta tcgcagcagt gagggtgagg ctgcccgcct cctcctccgt 300
 cgctgcccgc tctgggggga tgccacttgc ctccagctgg ccatgcaagc tgacgcccgt 360
 gccttctttg cccaggatgg ggtacagtct ctgctgacac agaagtgggt gggagatatg 420
 gccagcacta caccatcttg gcccttggtt ctgccttctt ttgcccctcc actcatctac 480
 accgcctca tcaccttcag gaaatcagaa gaggagccca cacgggagga gctagagttt 540
 gacatggata gtgtcattaa tggggaaggg cctgtcggga cggcggacc agccgagaag 600
 acgcccgtgg ggttcccgcg ccagtcgggc cgctccgggt gctgcggggg ccgctgcggg 660
 gggcgccggt gcctacgccg ctggttccac ttctggggcg cgccggtgac catcttcatg 720
 ggcaacgtgg tcagctacct gctgttctct ctgcttttct cgcggttgct gctcgtggat 780
 ttccagccgg cgccgcccgg ctccctggag ctgctgtctt atttctgggc ttacacgtg 840
 ctgtgcgagg aactgcgcca gggcctgagc ggagggggg gcagcctcgc cagcgggggc 900
 cccgggcctg gccatgcctc actgagccag cgctgcgcc tctacctcgc cgacagctgg 960
 aaccagtgcg acctagtggc tctcacctgc ttctcctg gcgtgggctg ccgctgacc 1020
 ccgggtttgt accacctggg ccgcactgtc ctctgcatcg acttcatggg ttacacgtg 1080
 cggtgtcttc acatcttcac ggtcaacaaa cagctggggc ccaagatcgt catcgtgagc 1140
 aagatgatga aggacgtgtt cttcttctct ttcttctcgc gcgtgtgggt ggtagcctat 1200
 ggcgtggcca cggagggggt cctgaggcca cgggacagt acttcccaag tatcctgcgc 1260
 cgctgttctt accgtcccta cctgcagatc ttccggcaga ttccccagga ggacatggac 1320
 gtggccctca tggagcacag caactgctcg tcggagcccg gcttctgggc acaccctcct 1380
 ggggcccagg cgggcacctg cgtctcccag tatgccaact ggctgggtgg gctgtctctc 1440
 gtcattcttc tgctcgtggc caacatctcg ctggtcaact tgctcattgc catgttcagt 1500
 tacacattcg gcaaagtaca gggcaacagc gatctctact ggaaggcgca gcgttaccgc 1560
 ctcatccggg aattccactc tcggcccgcg ctggcccgcg cctttatcgt catctcccac 1620
 ttgcgcctcc tgctcaggca attgtgcagg cgacccsnga gccccagcc gtcctcccgc 1680
 gccctcgagc atttccgggt ttacctttct aagggaagcg agcggaaagt gctaactggg 1740
 gaatcggtgc ataaggagaa ctttctgctg gcacgcgcta gggacaagcg ggagagcgac 1800
 tccgagmgtc tgaagcgcac gtcccagaag gtggacttgg cactgaaaca gctgggacac 1860

-37-

```

atccgcgagt acgaacagcg cctgaaagtg ctggagcggg aggtccagca gtgtacctcg 1920
gcccccgcac ctggtggcct tgtccttgag gtgagcccca tgtccatctg ggccactgtc 1980
aggaccacct ttgggagtgt catccttaca aaccacagca tgcccggtc ctcccagaac 2040
cagtcccagc ctgggaggat caaggcctgg atcccrggcc gttatccatc tggaggctgc 2100
agggtccttg gggtaacagg gaccacagac ccctcaccac tcacagattc ctcacactgg 2160
ggaaataaag ccatttcaga 2180

```

<210> 26
 <211> 725
 <212> PRT
 <213> Homo Sapiens

<220>
 <221> UNSURE
 <222> (553)...(553)
 <223> UNKNOWN

<221> UNSURE
 <222> (603)...(603)
 <223> UNKNOWN

<400> 26

Ser	Arg	Pro	Arg	Ile	Arg	His	Glu	Gly	Leu	Gly	Gln	Ala	Pro	Trp	Ser	1	5	10	15
Asp	Leu	Leu	Leu	Trp	Ala	Leu	Leu	Leu	Asn	Arg	Ala	Gln	Met	Ala	Met	20	25	30	
Tyr	Phe	Trp	Glu	Met	Gly	Ser	Asn	Ala	Val	Ser	Ser	Ala	Leu	Gly	Ala	35	40	45	
Cys	Leu	Leu	Leu	Arg	Val	Met	Ala	Arg	Leu	Glu	Pro	Asp	Ala	Glu	Glu	50	55	60	
Ala	Ala	Arg	Arg	Lys	Asp	Leu	Ala	Phe	Lys	Phe	Glu	Gly	Met	Gly	Val	65	70	75	80
Asp	Leu	Phe	Gly	Glu	Cys	Tyr	Arg	Ser	Ser	Glu	Val	Arg	Ala	Ala	Arg	85	90	95	
Leu	Leu	Leu	Arg	Arg	Cys	Pro	Leu	Trp	Gly	Asp	Ala	Thr	Cys	Leu	Gln	100	105	110	
Leu	Ala	Met	Gln	Ala	Asp	Ala	Arg	Ala	Phe	Phe	Ala	Gln	Asp	Gly	Val	115	120	125	
Gln	Ser	Leu	Leu	Thr	Gln	Lys	Trp	Trp	Gly	Asp	Met	Ala	Ser	Thr	Thr	130	135	140	
Pro	Ile	Trp	Ala	Leu	Val	Leu	Ala	Phe	Phe	Cys	Pro	Pro	Leu	Ile	Tyr	145	150	155	160
Thr	Arg	Leu	Ile	Thr	Phe	Arg	Lys	Ser	Glu	Glu	Pro	Thr	Arg	Glu		165	170	175	
Glu	Leu	Glu	Phe	Asp	Met	Asp	Ser	Val	Ile	Asn	Gly	Glu	Gly	Pro	Val	180	185	190	
Gly	Thr	Ala	Asp	Pro	Ala	Glu	Lys	Thr	Pro	Leu	Gly	Val	Pro	Arg	Gln	195	200	205	
Ser	Gly	Arg	Pro	Gly	Cys	Cys	Gly	Gly	Arg	Cys	Gly	Gly	Arg	Arg	Cys	210	215	220	
Leu	Arg	Arg	Trp	Phe	His	Phe	Trp	Gly	Ala	Pro	Val	Thr	Ile	Phe	Met	225	230	235	240
Gly	Asn	Val	Val	Ser	Tyr	Leu	Leu	Phe	Leu	Leu	Leu	Phe	Ser	Arg	Val	245	250	255	
Leu	Leu	Val	Asp	Phe	Gln	Pro	Ala	Pro	Pro	Gly	Ser	Leu	Glu	Leu	Leu	260	265	270	
Leu	Tyr	Phe	Trp	Ala	Phe	Thr	Leu	Leu	Cys	Glu	Glu	Leu	Arg	Gln	Gly	275	280	285	
Leu	Ser	Gly	Gly	Gly	Gly	Ser	Leu	Ala	Ser	Gly	Gly	Pro	Gly	Pro	Gly	290	295	300	

-38-

His Ala Ser Leu Ser Gln Arg Leu Arg Leu Tyr Leu Ala Asp Ser Trp
 305 310 315 320
 Asn Gln Cys Asp Leu Val Ala Leu Thr Cys Phe Leu Leu Gly Val Gly
 325 330 335
 Cys Arg Leu Thr Pro Gly Leu Tyr His Leu Gly Arg Thr Val Leu Cys
 340 345 350
 Ile Asp Phe Met Val Phe Thr Val Arg Leu Leu His Ile Phe Thr Val
 355 360 365
 Asn Lys Gln Leu Gly Pro Lys Ile Val Ile Val Ser Lys Met Met Lys
 370 375 380
 Asp Val Phe Phe Phe Leu Phe Phe Leu Gly Val Trp Leu Val Ala Tyr
 385 390 395 400
 Gly Val Ala Thr Glu Gly Leu Leu Arg Pro Arg Asp Ser Asp Phe Pro
 405 410 415
 Ser Ile Leu Arg Arg Val Phe Tyr Arg Pro Tyr Leu Gln Ile Phe Gly
 420 425 430
 Gln Ile Pro Gln Glu Asp Met Asp Val Ala Leu Met Glu His Ser Asn
 435 440 445
 Cys Ser Ser Glu Pro Gly Phe Trp Ala His Pro Pro Gly Ala Gln Ala
 450 455 460
 Gly Thr Cys Val Ser Gln Tyr Ala Asn Trp Leu Val Val Leu Leu Leu
 465 470 475 480
 Val Ile Phe Leu Leu Val Ala Asn Ile Leu Leu Val Asn Leu Leu Ile
 485 490 495
 Ala Met Phe Ser Tyr Thr Phe Gly Lys Val Gln Gly Asn Ser Asp Leu
 500 505 510
 Tyr Trp Lys Ala Gln Arg Tyr Arg Leu Ile Arg Glu Phe His Ser Arg
 515 520 525
 Pro Ala Leu Ala Pro Pro Phe Ile Val Ile Ser His Leu Arg Leu Leu
 530 535 540
 Leu Arg Gln Leu Cys Arg Arg Pro Xaa Ser Pro Gln Pro Ser Ser Pro
 545 550 555 560
 Ala Leu Glu His Phe Arg Val Tyr Leu Ser Lys Glu Ala Glu Arg Lys
 565 570 575
 Leu Leu Thr Trp Glu Ser Val His Lys Glu Asn Phe Leu Leu Ala Arg
 580 585 590
 Ala Arg Asp Lys Arg Glu Ser Asp Ser Glu Xaa Leu Lys Arg Thr Ser
 595 600 605
 Gln Lys Val Asp Leu Ala Leu Lys Gln Leu Gly His Ile Arg Glu Tyr
 610 615 620
 Glu Gln Arg Leu Lys Val Leu Glu Arg Glu Val Gln Gln Cys Thr Ser
 625 630 635 640
 Ala Pro Ala Pro Gly Gly Leu Val Leu Glu Val Ser Pro Met Ser Ile
 645 650 655
 Trp Ala Thr Val Arg Thr Thr Phe Gly Ser Val Ile Leu Thr Asn His
 660 665 670
 Ser Met Pro Gly Ser Ser Gln Asn Gln Ser Gln Pro Gly Arg Ile Lys
 675 680 685
 Ala Trp Ile Pro Gly Arg Tyr Pro Ser Gly Gly Cys Arg Val Leu Gly
 690 695 700
 Val Thr Gly Thr Thr Asp Pro Ser Pro Leu Thr Asp Ser Ser His Trp
 705 710 715 720
 Gly Asn Lys Ala Ile
 725

<210> 27
 <211> 7419
 <212> DNA
 <213> Homo Sapiens

<400> 27

cggggaccga	tccagcctcc	ggactctagc	ctaggctttt	gcaaaaagct	atthaggtga	60
cactatagaa	ggtacgcctg	caggtaccgg	tccggaattc	ccgggtcgac	ccacgcgtcc	120
gcagccccgt	cgccggcgga	ggcgggcgcg	ggcgcgtnc	ctgtggccag	tcacccggag	180
gagttggtcg	cacaattatg	aaagactcgg	cttctgctgc	tagcgccgga	gctgagttag	240
ttctgagaag	gtttccctgg	gcgttccttg	tccggcgggc	tctgctgccg	cctccggaga	300
cgcttcccga	tagatggcta	caggcccgcg	aggaggagga	ggtggagttg	ctgcccttcc	360
ggagtccgcc	ccgtgaggag	aatgtcccag	aaatcctgga	tagaaaagcac	tttgaccaag	420
agggaatgtg	tatatattat	accaagttcc	aaggaccctc	acagatgcct	tccaggatgt	480
caaatttgtc	agcaactcgt	caggtgtttt	tgtggtcgct	tggatcaagca	acatgcttgt	540
tttactgcaa	gtcttgccat	gaaatactca	gatgtgaaat	tgggtgacca	ttttaatcag	600
gcaatagaag	aatggtctgt	ggaaaagcat	acagaacaga	gccaacgga	tgcttatgga	660
gtcataaatt	ttcaaggggg	ttctcattcc	tacagagcta	agtatgtgag	gctatcatat	720
gacaccaaac	ctgaagtcac	tctgcaactt	ctgcttaaa	aatggcaaat	ggagttaccc	780
aaacttggtt	tctctgtaca	tgggggcatg	cagaaatttg	agcttcaccc	acgaatcaag	840
cagttgcttg	gaaaaggtct	tattaaagct	gcagttacaa	ctggagcctg	gattttaact	900
ggaggagtaa	acacaggtgt	ggcaaaacat	gttgagatg	ccctcaaaga	acatgcttcc	960
agatcatctc	gaaagatttg	cactatcgga	atagctccat	ggggagtgat	tgaaaacaga	1020
aatgatcttg	ttgggagaga	tgtggttgct	ccttatcaaa	ccttattgaa	ccccctgagc	1080
aaattgaatg	ttttgaataa	tctgcattcc	catttcata	tggatgatga	tggcactgtt	1140
ggaaagtatg	ggggggaagt	cagactgaga	agagaacttg	aaaaaactat	taatcagcaa	1200
agaattcatg	ctcaggattg	ccagggtgtc	cctgtggtgg	cacttatatt	tgagggtggg	1260
ccaaatgtta	tcttcacagt	tcttgaatac	cttcaggaaa	gccccctgt	tccagtagtt	1320
gtgtgtgaag	gaacaggcag	agctgcagat	ctgctagcgt	atattcataa	acaaacagaa	1380
gaaggaggga	atcttctctga	tgacagcagag	cccgatatta	ttccactat	caaaaaaaca	1440
tttaactttg	gccagaatga	agcacttcat	ttatttcaaa	cactgatgga	gtgcatgaaa	1500
agaaaggagc	ttatcactgt	tttccatatt	gggtcagatg	aacatcaaga	tatagatgta	1560
gcaataactta	ctgcactgct	aaaaggtact	aatgcactct	catttgacca	gcttatcctt	1620
acattggcat	gggatagagt	tgacattgcc	aaaaatcatg	tatttgttta	tgacagcag	1680
tggctgggtg	gacacttgga	acaagctatg	cttgatgctc	ttgtaatgga	tagagttgca	1740
tttgtaaaac	ttcttattga	aaatggagta	agcatgcata	aattccttac	cattccgaga	1800
ctggaagaac	tttacaacac	taaacaaggt	ccaactaatc	caatgctgtt	tcatcttggt	1860
cgagacgtca	aacaggga	tcttcctcca	ggatataaga	tcaactctgat	tgatatagga	1920
cttggttattg	aatatctcat	gggaggaacc	tacagatgca	cctatactag	gaaacgtttt	1980
cgattaatat	ataatagtct	tgggtggaat	aatcggaggt	ctggccgaaa	tacctccagc	2040
agcactcctc	agttgcgaaa	gagtcatgaa	tcttttgcca	atagggcaga	taaaaaggaa	2100
aaaatgaggc	ataaccattt	cattaagaca	gcacagccct	tccgaccaa	gattgatata	2160
gttatggaag	agggaagaa	gaaaagaacc	aaagatgaaa	ttgtagacat	tgatgatcca	2220
gaaaccaagc	agcttccctta	tccacttaat	gaacttttaa	tttgggcttg	ccttatgaa	2280
aggcaggtca	tggcccgctt	tttatggcaa	catggtgaag	aatcaatggc	taaagcatta	2340
gttgccctgta	agatctatcg	ttcaatggca	tatgaagcaa	agcagagtga	cctggtagat	2400
gatacttcag	aagaactaaa	acagtattcc	aatgatattg	gtcagttggc	cgttgaatta	2460
ttagaacagt	ccttcagaca	agatgaaacc	atggctatga	aattgctcac	ttatgaactg	2520
aagaactgga	gtaattcaac	ctgccttaag	ttagcagttt	cttcaagact	tagacctttt	2580
gtagctcaca	cctgtacaca	aatgttggtt	tctgatattg	ggatgggaag	gctgaatatg	2640
aggaaaaatt	cctggtacaa	ggtcatacta	agcattttag	ttccacctgc	catattgctg	2700
ttagagtata	aaactaaggc	tgaaatgtcc	catatcccac	aatctcaaga	tgctcatcag	2760
atgacaatgg	atgacagcga	aaacaacttt	cagaacataa	cagaagagat	ccccatggaa	2820
gtgtttaaag	aagtacggat	tttggtatagt	aatgaaggaa	agaatgagat	ggagatacaa	2880
atgaaatcaa	aaaagcttcc	aattacgcga	aagttttatg	ccttttatca	tgaccaaat	2940
gtaaaattct	ggtttaacac	gttggcatat	ttaggatttc	tgatgcttta	tacatttggt	3000
gttcttgtag	aaatggaaca	gttaccttca	gttcaagaat	ggattgttat	tgcttatatt	3060
tttacttatg	ccattgagaa	agtcctgtag	atctttatgt	ctgaagctgg	gaaagtaaac	3120
cagaagatta	aagtatggtt	tagtgattac	ttcaacatca	gtgatacaat	tgccataatt	3180
tctttcttca	ttggatttgg	actaagattt	ggagcaaaa	ggaactttgc	aaatgcatat	3240
gataactcatg	tttttgtggc	tggaagatta	atttactgtc	taaacataat	attttggtat	3300
gtgcgtttgc	tagattttct	agctgtaaa	caacaggcag	gaccttatgt	aatgatgatt	3360
ggaaaaatgg	tggccaatat	gttctacatt	gtagtgatta	tggctcttgt	attacttagt	3420
tttggtgttc	ccagaaaggc	aatactttat	cctcatgaag	caccatcttg	gactcttgct	3480
aaagatatag	tttttcaccc	atactggatg	atttttggtg	aagtttatgc	atacgaaatt	3540

gatgtgtgtg	caaatgattc	tgttatccct	caaactctgtg	gtcctggggac	gtgggtgact	3600
ccatttcttc	aagcagtccta	cctcttttgta	cagtatatca	ttatgggttaa	tccttcttatt	3660
gcattttttca	acaatgtgta	tttacaagtg	aaggcaattt	ccaatattgt	atggaagtac	3720
cagcgttatc	attttattat	ggcttatcat	gagaaaccag	ttctgcctcc	tccacttatac	3780
attcttagcc	atatagtttc	tctgttttgc	tgcatatgta	agagaagaaa	gaaagataag	3840
acttccgatg	gaccaaact	tttcttaaca	gaagaagatc	aaaagaaact	tcattgatttt	3900
gaagagcagt	gtgttgaaat	gtatttcaat	gaaaaagatg	acaaatttca	ttctgggagt	3960
gaagagagaa	ttcgtgtcac	ttttgaaaga	gtggaacaga	tgtgcattca	gattaaagaa	4020
gttggaatac	cacttagcta	cataaaaaga	tcattacaat	cattagattc	tcaaattggc	4080
catttgcaag	atctttcagc	cctgacggta	gatacattaa	aaacactcac	tgcccagaaa	4140
gcgtcggaag	ctagcaaaagt	tcataatgaa	atcacacgag	aactgagcat	ttccaaacac	4200
ttggctcaaa	accttattga	tgatggctct	gtaagacctt	ctgtatggaa	aaagcatggt	4260
gattggaata	cacttagctc	ctctcttcc	caagggtgatc	ttgaaagtaa	taatcctttt	4320
cattgtaata	ttttaatgaa	agatgacaaa	gatccccagt	gtaatatatt	tggtcaagac	4380
ttacctgcag	tacccagag	aaaagaattt	aattttccag	aggctgggtc	ctcttctggt	4440
gccttatttc	caagtgtctg	ttcccctcca	gaactgcgac	agagactaca	tggggtagaa	4500
ctcttaaaaa	tatttaataa	aaatcaaaaa	tgtaggcatt	catctactag	cataccacat	4560
ctgtctatccc	caccaaccaa	attttttgtt	agtacacccat	ctcagccaa	ttgcaaaagc	4620
cacttggaata	ctggaaccaa	agatcaagaa	actgtttgct	ctaaagctac	agaaggagat	4680
aatacagaat	ttggagcatt	tgtaggacac	agagatagca	tggatttaca	gaggtttaaa	4740
gaaacatcaa	acaagataaa	aatactatcc	aataacaata	cttctgaaaa	cactttgaaa	4800
cgagtgtggt	ctcttgctgg	atttactgac	tgtcacagaa	cttccattcc	tgttcattca	4860
aaacaagcag	aaaaaatcag	tagaaggcca	tctaccgaag	acactcatga	agtagattcc	4920
aaagcagctt	taataccgga	ttgggttaca	gatagaccat	caaacagaga	aatgccatct	4980
gaagaaggaa	cattaaatgg	tctcacttct	ccattttaagc	cagctatgga	tacaaattac	5040
tattattcag	ctgtggaaa	aaataacttg	atgaggttat	cacagagcat	tccatttaca	5100
cctgtgcctc	caagagggga	gcctgtcaca	gtgtatcggt	tggaaagag	ttcacccaac	5160
atactaaata	acagcatgtc	ttcttggtca	caactaggcc	tctgtgcca	aatagagttt	5220
ttaagcaaa	aggagatggg	aggaggttta	cgaagagctg	tcaaagtaca	gtgtacctgg	5280
tcagaacatg	atatcctcaa	atcagggcac	ctttatatta	tcaaactctt	tcttccagag	5340
gtgggttaata	cttggtcaag	tatttataaa	gaagatacac	ttctgcattc	ctgtctgaga	5400
gaaattcaac	aacagagagc	agcacaaaag	cttacgtttg	cctttaatca	aatgaaacct	5460
aaatccatac	catattctcc	aagggttcctt	gaagttttcc	tgctgtattg	ccattcagca	5520
ggacagtggt	ttgctgtgga	agaatgtatg	actggagaat	ttagaaaaata	caacaataat	5580
aatggagatg	agattattcc	aactaatact	ctggaagaga	tcattgctagc	ctttagccac	5640
tggacttacg	aatatacaag	aggggagtta	ctgggtactg	atttgcaagg	tggttggtgaa	5700
aatttgactg	acccatctgt	gataaaagca	gaagaaaaga	gatcctgtga	tatgggtttt	5760
ggcccagcaa	atctaggaga	agatgcaatt	aaaaacttca	gagcaaaaaca	tcactgtaat	5820
tcttgctgtc	gaaagcttaa	acttcagat	ctgaagagga	atgattatac	gcctgataaa	5880
attatatttc	ctcaggatga	gccttcagat	ttgaattctc	agcctggaaa	ttccaccaaa	5940
gaatcagaat	caactaattc	tgcttctgtc	atgttataat	attaatatta	ctgaatcatt	6000
ggttttgcct	gcacctcaca	gaaatgttac	tgtgtcactt	ttccctcggg	aggaaattgt	6060
ttggtaatat	agaaagggtg	atgcaagtgg	aatttgctga	ctccagcaca	gttaaaagg	6120
caatattctt	ttgacctgat	taatcagtca	gaaagtcctt	ataggataga	gctggcagct	6180
gagaaatttt	aaaggtaatt	gataattagt	atlttgtaact	ttttaagggt	ctctttgtat	6240
agcagaggat	ctcatttgac	tttggtttga	tgagggtgat	gccctctctt	atgtggtaca	6300
ataccattaa	ccaaaggtag	gtgtccatgc	agattttatt	ggcagctggt	ttattgccat	6360
tcaactaggg	aaatgaagaa	atcacgcagc	cttttggtta	aatggcagtc	aaaaattttcc	6420
tcagtgtatt	tagtgtgttc	agtgtatgata	tcactgggtc	ccaactagat	gcttggtggc	6480
cacgggaagg	gaaatgactt	gttctaattc	taggttcaca	gaggtatgag	aagcctgaac	6540
tgaagaccat	tttcaagagg	gacggtattt	atgaatcagg	gttaggctcc	atattttaaag	6600
atagagccag	tttttttttt	aaatagaacc	caaattgtgt	aaaaatgtta	attgggtttt	6660
ttaaacattg	tttttatcaag	tcactgttaa	gtagaagaaa	gccatggtaa	actgatcat	6720
aacctaaatt	ataaaagcag	aaaccttaact	cactcgtcaa	gggaagtta	cttttgagga	6780
aagttaaagt	acttttttcc	ctatctgtat	ctatagcaac	aaaccagaa	ttacaaactt	6840
ctccaaagat	tttattgatt	gttatatcaa	atcagaatgt	aaacatgaac	tcttgcatat	6900
atttaaaatt	gtgttggaac	atttgaacat	gaatgctgtt	tgggtactta	agaaatttrat	6960
tcagtnngat	tatcattatg	tganactggc	agattgcagt	gcanccttat	gccaataaaa	7020
tgtaatttar	cagccccaga	tattgttgaa	tattcaacaa	taacaagaaa	agcttttcat	7080
ctaagtttta	tgctttaatt	ttttttcttt	ttttttcttt	ttcttttgtt	tccttggtac	7140

-41-

taattttaat	ttttatttgg	aagggagcag	tataaagctt	atttgtattt	agtagtgtat	7200
ctcatagata	cagacaaggc	aagagatgat	aagctgttta	aatagtgttt	aatattgatt	7260
gggggtgggg	agaaagaaaa	agtgtattac	ttaaagatac	tatatacskt	ttktatatca	7320
ttaaatcttt	aaaagaaatn	naataaattt	attgtttnc	aaaaaaaaac	ccnntaaaaa	7380
aaaaagggcg	gccccctctag	aggatccctc	gaggggccc			7419

<210> 28
 <211> 1865
 <212> PRT
 <213> Homo Sapiens

<400> 28

Met	Ser	Gln	Lys	Ser	Trp	Ile	Glu	Ser	Thr	Leu	Thr	Lys	Arg	Glu	Cys
1			5						10					15	
Val	Tyr	Ile	Ile	Pro	Ser	Ser	Lys	Asp	Pro	His	Arg	Cys	Leu	Pro	Gly
			20					25					30		
Cys	Gln	Ile	Cys	Gln	Gln	Leu	Val	Arg	Cys	Phe	Cys	Gly	Arg	Leu	Val
		35					40					45			
Lys	Gln	His	Ala	Cys	Phe	Thr	Ala	Ser	Leu	Ala	Met	Lys	Tyr	Ser	Asp
		50				55					60				
Val	Lys	Leu	Gly	Asp	His	Phe	Asn	Gln	Ala	Ile	Glu	Glu	Trp	Ser	Val
65					70				75					80	
Glu	Lys	His	Thr	Glu	Gln	Ser	Pro	Thr	Asp	Ala	Tyr	Gly	Val	Ile	Asn
				85					90					95	
Phe	Gln	Gly	Gly	Ser	His	Ser	Tyr	Arg	Ala	Lys	Tyr	Val	Arg	Leu	Ser
			100					105					110		
Tyr	Asp	Thr	Lys	Pro	Glu	Val	Ile	Leu	Gln	Leu	Leu	Leu	Lys	Glu	Trp
			115				120					125			
Gln	Met	Glu	Leu	Pro	Lys	Leu	Val	Ile	Ser	Val	His	Gly	Gly	Met	Gln
			130				135				140				
Lys	Phe	Glu	Leu	His	Pro	Arg	Ile	Lys	Gln	Leu	Gly	Lys	Gly	Leu	
145					150					155				160	
Ile	Lys	Ala	Ala	Val	Thr	Thr	Gly	Ala	Trp	Ile	Leu	Thr	Gly	Gly	Val
				165					170					175	
Asn	Thr	Gly	Val	Ala	Lys	His	Val	Gly	Asp	Ala	Leu	Lys	Glu	His	Ala
			180					185					190		
Ser	Arg	Ser	Ser	Arg	Lys	Ile	Cys	Thr	Ile	Gly	Ile	Ala	Pro	Trp	Gly
		195					200					205			
Val	Ile	Glu	Asn	Arg	Asn	Asp	Leu	Val	Gly	Arg	Asp	Val	Val	Ala	Pro
		210				215					220				
Tyr	Gln	Thr	Leu	Leu	Asn	Pro	Leu	Ser	Lys	Leu	Asn	Val	Leu	Asn	Asn
225					230					235				240	
Leu	His	Ser	His	Phe	Ile	Leu	Val	Asp	Asp	Gly	Thr	Val	Gly	Lys	Tyr
			245						250					255	
Gly	Ala	Glu	Val	Arg	Leu	Arg	Arg	Glu	Leu	Glu	Lys	Thr	Ile	Asn	Gln
			260					265					270		
Gln	Arg	Ile	His	Ala	Arg	Ile	Gly	Gln	Gly	Val	Pro	Val	Val	Ala	Leu
			275				280					285			
Ile	Phe	Glu	Gly	Gly	Pro	Asn	Val	Ile	Leu	Thr	Val	Leu	Glu	Tyr	Leu
		290				295					300				
Gln	Glu	Ser	Pro	Pro	Val	Pro	Val	Val	Val	Cys	Glu	Gly	Thr	Gly	Arg
305					310					315				320	
Ala	Ala	Asp	Leu	Leu	Ala	Tyr	Ile	His	Lys	Gln	Thr	Glu	Glu	Gly	Gly
			325						330					335	
Asn	Leu	Pro	Asp	Ala	Ala	Glu	Pro	Asp	Ile	Ile	Ser	Thr	Ile	Lys	Lys
			340					345					350		
Thr	Phe	Asn	Phe	Gly	Gln	Asn	Glu	Ala	Leu	His	Leu	Phe	Gln	Thr	Leu
		355					360						365		

-42-

Met	Glu	Cys	Met	Lys	Arg	Lys	Glu	Leu	Ile	Thr	Val	Phe	His	Ile	Gly
	370					375					380				
Ser	Asp	Glu	His	Gln	Asp	Ile	Asp	Val	Ala	Ile	Leu	Thr	Ala	Leu	Leu
385					390					395					400
Lys	Gly	Thr	Asn	Ala	Ser	Ala	Phe	Asp	Gln	Leu	Ile	Leu	Thr	Leu	Ala
			405					410							415
Trp	Asp	Arg	Val	Asp	Ile	Ala	Lys	Asn	His	Val	Phe	Val	Tyr	Gly	Gln
	420							425					430		
Gln	Trp	Leu	Val	Gly	Ser	Leu	Glu	Gln	Ala	Met	Leu	Asp	Ala	Leu	Val
	435						440					445			
Met	Asp	Arg	Val	Ala	Phe	Val	Lys	Leu	Leu	Ile	Glu	Asn	Gly	Val	Ser
450						455					460				
Met	His	Lys	Phe	Leu	Thr	Ile	Pro	Arg	Leu	Glu	Glu	Leu	Tyr	Asn	Thr
465					470					475					480
Lys	Gln	Gly	Pro	Thr	Asn	Pro	Met	Leu	Phe	His	Leu	Val	Arg	Asp	Val
			485						490					495	
Lys	Gln	Gly	Asn	Leu	Pro	Pro	Gly	Tyr	Lys	Ile	Thr	Leu	Ile	Asp	Ile
			500					505					510		
Gly	Leu	Val	Ile	Glu	Tyr	Leu	Met	Gly	Gly	Thr	Tyr	Arg	Cys	Thr	Tyr
	515						520					525			
Thr	Arg	Lys	Arg	Phe	Arg	Leu	Ile	Tyr	Asn	Ser	Leu	Gly	Gly	Asn	Asn
530						535					540				
Arg	Arg	Ser	Gly	Arg	Asn	Thr	Ser	Ser	Ser	Thr	Pro	Gln	Leu	Arg	Lys
545					550					555					560
Ser	His	Glu	Ser	Phe	Gly	Asn	Arg	Ala	Asp	Lys	Lys	Glu	Lys	Met	Arg
			565						570					575	
His	Asn	His	Phe	Ile	Lys	Thr	Ala	Gln	Pro	Phe	Arg	Pro	Lys	Ile	Asp
			580					585					590		
Thr	Val	Met	Glu	Glu	Gly	Lys	Lys	Lys	Arg	Thr	Lys	Asp	Glu	Ile	Val
	595					600						605			
Asp	Ile	Asp	Asp	Pro	Glu	Thr	Lys	Arg	Phe	Pro	Tyr	Pro	Leu	Asn	Glu
610						615					620				
Leu	Leu	Ile	Trp	Ala	Cys	Leu	Met	Lys	Arg	Gln	Val	Met	Ala	Arg	Phe
625					630					635					640
Leu	Trp	Gln	His	Gly	Glu	Glu	Ser	Met	Ala	Lys	Ala	Leu	Val	Ala	Cys
			645						650					655	
Lys	Ile	Tyr	Arg	Ser	Met	Ala	Tyr	Glu	Ala	Lys	Gln	Ser	Asp	Leu	Val
			660					665					670		
Asp	Asp	Thr	Ser	Glu	Glu	Leu	Lys	Gln	Tyr	Ser	Asn	Asp	Phe	Gly	Gln
	675						680					685			
Leu	Ala	Val	Glu	Leu	Leu	Glu	Gln	Ser	Phe	Arg	Gln	Asp	Glu	Thr	Met
690						695					700				
Ala	Met	Lys	Leu	Leu	Thr	Tyr	Glu	Leu	Lys	Asn	Trp	Ser	Asn	Ser	Thr
705					710					715					720
Cys	Leu	Lys	Leu	Ala	Val	Ser	Ser	Arg	Leu	Arg	Pro	Phe	Val	Ala	His
			725						730					735	
Thr	Cys	Thr	Gln	Met	Leu	Leu	Ser	Asp	Met	Trp	Met	Gly	Arg	Leu	Asn
			740					745				750			
Met	Arg	Lys	Asn	Ser	Trp	Tyr	Lys	Val	Ile	Leu	Ser	Ile	Leu	Val	Pro
	755						760					765			
Pro	Ala	Ile	Leu	Leu	Leu	Glu	Tyr	Lys	Thr	Lys	Ala	Glu	Met	Ser	His
770						775					780				
Ile	Pro	Gln	Ser	Gln	Asp	Ala	His	Gln	Met	Thr	Met	Asp	Asp	Ser	Glu
785					790					795					800
Asn	Asn	Phe	Gln	Asn	Ile	Thr	Glu	Glu	Ile	Pro	Met	Glu	Val	Phe	Lys
			805						810					815	
Glu	Val	Arg	Ile	Leu	Asp	Ser	Asn	Glu	Gly	Lys	Asn	Glu	Met	Glu	Ile
			820					825				830			
Gln	Met	Lys	Ser	Lys	Lys	Leu	Pro	Ile	Thr	Arg	Lys	Phe	Tyr	Ala	Phe
	835						840					845			

-43-

Tyr	His	Ala	Pro	Ile	Val	Lys	Phe	Trp	Phe	Asn	Thr	Leu	Ala	Tyr	Leu
850						855				860					
Gly	Phe	Leu	Met	Leu	Tyr	Thr	Phe	Val	Val	Leu	Val	Gln	Met	Glu	Gln
865					870					875					880
Leu	Pro	Ser	Val	Gln	Glu	Trp	Ile	Val	Ile	Ala	Tyr	Ile	Phe	Thr	Tyr
				885					890					895	
Ala	Ile	Glu	Lys	Val	Arg	Glu	Ile	Phe	Met	Ser	Glu	Ala	Gly	Lys	Val
			900					905					910		
Asn	Gln	Lys	Ile	Lys	Val	Trp	Phe	Ser	Asp	Tyr	Phe	Asn	Ile	Ser	Asp
		915					920					925			
Thr	Ile	Ala	Ile	Ile	Ser	Phe	Phe	Ile	Gly	Phe	Gly	Leu	Arg	Phe	Gly
	930					935					940				
Ala	Lys	Trp	Asn	Phe	Ala	Asn	Ala	Tyr	Asp	Asn	His	Val	Phe	Val	Ala
945				950						955					960
Gly	Arg	Leu	Ile	Tyr	Cys	Leu	Asn	Ile	Ile	Phe	Trp	Tyr	Val	Arg	Leu
			965						970					975	
Leu	Asp	Phe	Leu	Ala	Val	Asn	Gln	Gln	Ala	Gly	Pro	Tyr	Val	Met	Met
			980					985					990		
Ile	Gly	Lys	Met	Val	Ala	Asn	Met	Phe	Tyr	Ile	Val	Val	Ile	Met	Ala
	995					1000						1005			
Leu	Val	Leu	Leu	Ser	Phe	Gly	Val	Pro	Arg	Lys	Ala	Ile	Leu	Tyr	Pro
	1010					1015					1020				
His	Glu	Ala	Pro	Ser	Trp	Thr	Leu	Ala	Lys	Asp	Ile	Val	Phe	His	Pro
1025				1030						1035					1040
Tyr	Trp	Met	Ile	Phe	Gly	Glu	Val	Tyr	Ala	Tyr	Glu	Ile	Asp	Val	Cys
			1045						1050					1055	
Ala	Asn	Asp	Ser	Val	Ile	Pro	Gln	Ile	Cys	Gly	Pro	Gly	Thr	Trp	Leu
			1060					1065					1070		
Thr	Pro	Phe	Leu	Gln	Ala	Val	Tyr	Leu	Phe	Val	Gln	Tyr	Ile	Ile	Met
	1075						1080					1085			
Val	Asn	Leu	Leu	Ile	Ala	Phe	Phe	Asn	Asn	Val	Tyr	Leu	Gln	Val	Lys
	1090					1095					1100				
Ala	Ile	Ser	Asn	Ile	Val	Trp	Lys	Tyr	Gln	Arg	Tyr	His	Phe	Ile	Met
1105			1110							1115					1120
Ala	Tyr	His	Glu	Lys	Pro	Val	Leu	Pro	Pro	Pro	Leu	Ile	Ile	Leu	Ser
			1125						1130					1135	
His	Ile	Val	Ser	Leu	Phe	Cys	Cys	Ile	Cys	Lys	Arg	Arg	Lys	Lys	Asp
	1140							1145					1150		
Lys	Thr	Ser	Asp	Gly	Pro	Lys	Leu	Phe	Leu	Thr	Glu	Glu	Asp	Gln	Lys
	1155						1160						1165		
Lys	Leu	His	Asp	Phe	Glu	Glu	Gln	Cys	Val	Glu	Met	Tyr	Phe	Asn	Glu
	1170					1175					1180				
Lys	Asp	Asp	Lys	Phe	His	Ser	Gly	Ser	Glu	Glu	Arg	Ile	Arg	Val	Thr
1185			1190							1195					1200
Phe	Glu	Arg	Val	Glu	Gln	Met	Cys	Ile	Gln	Ile	Lys	Glu	Val	Gly	Asp
			1205						1210					1215	
Arg	Val	Asn	Tyr	Ile	Lys	Arg	Ser	Leu	Gln	Ser	Leu	Asp	Ser	Gln	Ile
			1220					1225					1230		
Gly	His	Leu	Gln	Asp	Leu	Ser	Ala	Leu	Thr	Val	Asp	Thr	Leu	Lys	Thr
	1235						1240					1245			
Leu	Thr	Ala	Gln	Lys	Ala	Ser	Glu	Ala	Ser	Lys	Val	His	Asn	Glu	Ile
	1250					1255					1260				
Thr	Arg	Glu	Leu	Ser	Ile	Ser	Lys	His	Leu	Ala	Gln	Asn	Leu	Ile	Asp
1265			1270							1275					1280
Asp	Gly	Pro	Val	Arg	Pro	Ser	Val	Trp	Lys	Lys	His	Gly	Val	Val	Asn
			1285						1290					1295	
Thr	Leu	Ser	Ser	Ser	Leu	Pro	Gln	Gly	Asp	Leu	Glu	Ser	Asn	Asn	Pro
			1300					1305					1310		
Phe	His	Cys	Asn	Ile	Leu	Met	Lys	Asp	Asp	Lys	Asp	Pro	Gln	Cys	Asn
		1315					1320					1325			

-44-

Ile Phe Gly Gln Asp Leu Pro Ala Val Pro Gln Arg Lys Glu Phe Asn
 1330 1335 1340
 Phe Pro Glu Ala Gly Ser Ser Ser Gly Ala Leu Phe Pro Ser Ala Val
 1345 1350 1355 1360
 Ser Pro Pro Glu Leu Arg Gln Arg Leu His Gly Val Glu Leu Leu Lys
 1365 1370 1375
 Ile Phe Asn Lys Asn Gln Lys Leu Gly Ser Ser Ser Thr Ser Ile Pro
 1380 1385 1390
 His Leu Ser Ser Pro Pro Thr Lys Phe Phe Val Ser Thr Pro Ser Gln
 1395 1400 1405
 Pro Ser Cys Lys Ser His Leu Glu Thr Gly Thr Lys Asp Gln Glu Thr
 1410 1415 1420
 Val Cys Ser Lys Ala Thr Glu Gly Asp Asn Thr Glu Phe Gly Ala Phe
 1425 1430 1435 1440
 Val Gly His Arg Asp Ser Met Asp Leu Gln Arg Phe Lys Glu Thr Ser
 1445 1450 1455
 Asn Lys Ile Lys Ile Leu Ser Asn Asn Thr Ser Glu Asn Thr Leu
 1460 1465 1470
 Lys Arg Val Ser Ser Leu Ala Gly Phe Thr Asp Cys His Arg Thr Ser
 1475 1480 1485
 Ile Pro Val His Ser Lys Gln Ala Glu Lys Ile Ser Arg Arg Pro Ser
 1490 1495 1500
 Thr Glu Asp Thr His Glu Val Asp Ser Lys Ala Ala Leu Ile Pro Asp
 1505 1510 1515 1520
 Trp Leu Gln Asp Arg Pro Ser Asn Arg Glu Met Pro Ser Glu Glu Gly
 1525 1530 1535
 Thr Leu Asn Gly Leu Thr Ser Pro Phe Lys Pro Ala Met Asp Thr Asn
 1540 1545 1550
 Tyr Tyr Tyr Ser Ala Val Glu Arg Asn Asn Leu Met Arg Leu Ser Gln
 1555 1560 1565
 Ser Ile Pro Phe Thr Pro Val Pro Pro Arg Gly Glu Pro Val Thr Val
 1570 1575 1580
 Tyr Arg Leu Glu Glu Ser Ser Pro Asn Ile Leu Asn Asn Ser Met Ser
 1585 1590 1595 1600
 Ser Trp Ser Gln Leu Gly Leu Cys Ala Lys Ile Glu Phe Leu Ser Lys
 1605 1610 1615
 Glu Glu Met Gly Gly Gly Leu Arg Arg Ala Val Lys Val Gln Cys Thr
 1620 1625 1630
 Trp Ser Glu His Asp Ile Leu Lys Ser Gly His Leu Tyr Ile Ile Lys
 1635 1640 1645
 Ser Phe Leu Pro Glu Val Val Asn Thr Trp Ser Ser Ile Tyr Lys Glu
 1650 1655 1660
 Asp Thr Val Leu His Leu Cys Leu Arg Glu Ile Gln Gln Gln Arg Ala
 1665 1670 1675 1680
 Ala Gln Lys Leu Thr Phe Ala Phe Asn Gln Met Lys Pro Lys Ser Ile
 1685 1690 1695
 Pro Tyr Ser Pro Arg Phe Leu Glu Val Phe Leu Leu Tyr Cys His Ser
 1700 1705 1710
 Ala Gly Gln Trp Phe Ala Val Glu Glu Cys Met Thr Gly Glu Phe Arg
 1715 1720 1725
 Lys Tyr Asn Asn Asn Asn Gly Asp Glu Ile Ile Pro Thr Asn Thr Leu
 1730 1735 1740
 Glu Glu Ile Met Leu Ala Phe Ser His Trp Thr Tyr Glu Tyr Thr Arg
 1745 1750 1755 1760
 Gly Glu Leu Leu Val Leu Asp Leu Gln Gly Val Gly Glu Asn Leu Thr
 1765 1770 1775
 Asp Pro Ser Val Ile Lys Ala Glu Glu Lys Arg Ser Cys Asp Met Val
 1780 1785 1790
 Phe Gly Pro Ala Asn Leu Gly Glu Asp Ala Ile Lys Asn Phe Arg Ala
 1795 1800 1805

-45-

Lys His His Cys Asn Ser Cys Cys Arg Lys Leu Lys Leu Pro Asp Leu
 1810 1815 1820
 Lys Arg Asn Asp Tyr Thr Pro Asp Lys Ile Ile Phe Pro Gln Asp Glu
 1825 1830 1835 1840
 Pro Ser Asp Leu Asn Leu Gln Pro Gly Asn Ser Thr Lys Glu Ser Glu
 1845 1850 1855
 Ser Thr Asn Ser Val Arg Leu Met Leu
 1860 1865

<210> 29
 <211> 4061
 <212> DNA
 <213> Homo Sapiens

<400> 29
 ggtctggaag cagagccggc ggaggggagcg ccggggccct gggctgcagg aggttgccgc 60
 ggccgcggca gcatggtggt gccggagaag gagcagagct ggatcccca gatcttcaag 120
 aagaagacct gcacgacgtt catagttagc tccacagatc cgggaggagc cttgtgccaag 180
 tgtggggcgcc cccggaccgc ccaccccga gtggccatgg aggatgcctt cggggcagcc 240
 gtgtgacccg tgtgggacag cgatgcacac accacggaga agcccaccga tgcctacgga 300
 gagctggact tcacggggc cgcccgcaag cacagcaatt tcctccggct ctctgaccga 360
 acggatccag ctgcagttta tagtctggtc acacgcacat ggggcttccg tgcctcgaac 420
 ctggtggtgt cagtgcctggg gggatcgggg ggccccgtcc tccagacctg gctgcaggac 480
 ctgtgcgtc gtgggctggt gcgggctgcc cagagcacag gagcctggat tgcactggg 540
 ggtctgcaca cgggcatcgg ccggcatggt ggtgtggctg tacgggacca tcagatggcc 600
 agcactgggg gcaccaaggt ggtggccatg ggtgtggccc cctggggtgt ggtccggaat 660
 agagacaccc tcatcaaccc caagggtcgt ttccctgcga ggtaccggtg gcgcggtgac 720
 ccggaggacg ggggtccagtt tcccctggac tacaactact cggccttctt cctggtggac 780
 gacggcacac acggctgcct gggggggcag aaccgcttcc gcttgcgcct ggagtcctac 840
 atctcacagc agaagacggg cgtgggaggg actggaattg acatccctgt cctgctcctc 900
 ctgattgatg gtgatgagaa gatgttgacg cgaatagaga acgccacca ggctcagctc 960
 ccatgtctcc tcgtggctgg ctacggggga gctgcggaact gcctggcgga gacctggaa 1020
 gacactctgg cccagggag tgggggagcc aggcaaggcg aagcccaga tcgaatcagg 1080
 cgtttctttc ccaaaaggga ccttgaggtc ctgcaggccc aggtggagag gattatgacc 1140
 cggaaaggagc tcctgacagt ctattcttct gaggatgggt ctgaggaatt cgagaccata 1200
 gttttgaagg cccttgtaga ggcctgtggg agctcggagg cctcagccta cctggatgag 1260
 ctgcgttttg ctgtggttga gaaccgcgtg gacattgccc agagtgaact ctttcggggg 1320
 gacatccaat ggcggtcctt ccatctcgaa gcttccctca tggacgccct gctgaatgac 1380
 cggctcagtg tcgtgcgctt gctcatttcc caggccctca gcctgggcca cttcctgacc 1440
 ccgatgcgcc tggcccaact ctacagcgcg gcgcctcca actcgtcat ccgcaacctt 1500
 ttggaccagg cgtccacag cgacggcacc aaagcccag ccctaaaagg gggagctgcg 1560
 gagctccggc ccctgacgt ggggcatgtg ctgaggatgc tgctggggaa gatgtgcgcg 1620
 ccgaggtacc cctccggggg cgcctgggac cctcaccag gccagggtt cggggagagc 1680
 atgtatctgc tctcggacaa ggccacctcg ccgctctcgc tggatgctgg cctcgggcag 1740
 gccccttga gcgacctgct tctttgggca ctgttgctga acagggcaca gatggccatg 1800
 tacttctggg agatgggttc caatgcagtt tcctcagctc ttggggcctg tttgctgctc 1860
 cgggtgatgg cacgcctgga gcctgacgt gaggaggaag cacggaggaa agacctggcg 1920
 ttcaagtttg aggggatggg cgttgacctc tttggcgagt gctatcgag cagtgaggtg 1980
 agggctgccc gcctcctcct ccgtcgtcgc ccgctctggg gggatgccac ttgcctccag 2040
 ctggccatgc aagctgacgc ccgtgccttc tttgccagg atgggtaca gtctctgctg 2100
 acacagaagt ggtggggaga tatggccagc actacacca tctggggcct ggttctcgcc 2160
 ttcttttgcc ctccactcat ctacaccgc ctcacacct tcaggaaatc agaagaggag 2220
 cccacacggg aggagctaga gtttgacatg gatagtgtca ttaatgggga agggcctgtc 2280
 gggacggcgg accagccga gaagacggc ctgggggtcc cgcgccagtc gggccgtccg 2340
 ggttgctgcg ggggcccgtg cggggggcgc cgggtcctac gccgctggtt ccacttctgg 2400
 ggcgcgcggg tgaccatct catgggcaac gtggtcagct acctgctgt cctgctctt 2460
 ttctcgcggg tgcctcgtt ggatttccag ccggcgccgc ccggctcctt ggagctgctg 2520
 ctctatttct gggctttcac gctgctgtgc gaggaactgc gccaggccct gagcggaggc 2580
 gggggcagcc tcgccagcgg gggccccggg cctggccatg cctcactgag ccagcgcctg 2640
 cgcctctacc tcgccgacag ctggaaccag tgcgacctag tggtctctac ctgcttctc 2700

-46-

ctgggcggtg	gctgccggct	gaccccggtt	ttgtaccacc	tgggcccgcac	tgtcctctgc	2760
atcgacttca	tggtttttcac	ggtgcggtctg	cttcacatct	tcacgggtcaa	caaacagctg	2820
gggccaaga	tcgtcatcgt	gagcaagatg	atgaaggacg	tggtcttctt	cctcttcttc	2880
ctcggcggtg	ggctggtagc	ctatggcggtg	gccacggagg	ggctcctgag	gccacgggac	2940
agtgacttcc	caagtatcct	gcgccgcgtc	ttctaccgtc	cctacctgca	gatcttcggg	3000
cagattcccc	aggaggacat	ggacgtggcc	ctcatggagc	acagcaactg	ctcgtcggag	3060
cccggcttct	gggcacaccc	tcctggggcc	caggcgggca	cctgcgtctc	ccagtatgcc	3120
aactggctgg	tggtgctgct	cctcgtcatc	ttcctgctcg	tggccaacat	cctgctggtc	3180
aacttgctca	ttgccatggt	cagttacaca	ttcggcaaaag	tacagggcaa	cagcgatctc	3240
tactggaagg	cgcagcggtt	ccgcctcatc	cgggaattcc	actctcggcc	cgcgtgggcc	3300
ccgcccttta	tcgtcatctc	ccacttgctc	ctcctgctca	ggcaattgtg	caggcgaccc	3360
cggagccccc	agccgtcctc	cccggccctc	gagcatttcc	gggtttacct	ttctaaggaa	3420
gccgagcgga	agctgctaac	gtgggaatcg	gtgcataaag	agaactttct	gctggcacgc	3480
gctagggaca	agcgggagag	cgactccgag	cgtctgaagc	gcacgtccca	gaaggtggac	3540
ttggcactga	aacagctggg	acacatccgc	gagtacgaac	agcgccctgaa	agtgtggag	3600
cgggaggtcc	agcagtgtag	ccgcgtcctg	gggtgggtgg	ccgaggccct	gagccgtctc	3660
gccttgctgc	ccccagggtg	gccgccaccc	cctgacctgc	ctgggtccaa	agactgagcc	3720
ctgctggcgg	acttcaagga	gaagccccc	cagggggattt	tgctcctaga	gtaaggctca	3780
tctgggcctc	ggcccccgca	cctgggtggcc	ttgtccttga	ggtgagcccc	atgtccatct	3840
gggccactgt	caggaccacc	tttgggagtg	tcatccttac	aaaccacagc	atgcccggt	3900
cctcccagaa	ccagtcccag	cctgggagga	tcaaggcctg	gatccccggg	cgttatccat	3960
ctggaggtcg	cagggtcctt	ggggtaacag	gcaccacaga	cccctcacca	ctcacagatt	4020
cctcacactg	gggaaataaa	gccatttcag	aggaaaaaaa	a		4061

<210> 30

<211> 1214

<212> PRT

<213> Homo Sapiens

<400> 30

Met	Val	Val	Pro	Glu	Lys	Glu	Gln	Ser	Trp	Ile	Pro	Lys	Ile	Phe	Lys
1				5					10					15	
Lys	Lys	Thr	Cys	Thr	Thr	Phe	Ile	Val	Asp	Ser	Thr	Asp	Pro	Gly	Gly
			20					25					30		
Thr	Leu	Cys	Gln	Cys	Gly	Arg	Pro	Arg	Thr	Ala	His	Pro	Ala	Val	Ala
		35					40					45			
Met	Glu	Asp	Ala	Phe	Gly	Ala	Ala	Val	Val	Thr	Val	Trp	Asp	Ser	Asp
	50					55					60				
Ala	His	Thr	Thr	Glu	Lys	Pro	Thr	Asp	Ala	Tyr	Gly	Glu	Leu	Asp	Phe
	65				70					75				80	
Thr	Gly	Ala	Gly	Arg	Lys	His	Ser	Asn	Phe	Leu	Arg	Leu	Ser	Asp	Arg
				85				90						95	
Thr	Asp	Pro	Ala	Ala	Val	Tyr	Ser	Leu	Val	Thr	Arg	Thr	Trp	Gly	Phe
			100					105					110		
Arg	Ala	Pro	Asn	Leu	Val	Val	Ser	Val	Leu	Gly	Gly	Ser	Gly	Gly	Pro
		115					120					125			
Val	Leu	Gln	Thr	Trp	Leu	Gln	Asp	Leu	Leu	Arg	Arg	Gly	Leu	Val	Arg
	130					135					140				
Ala	Ala	Gln	Ser	Thr	Gly	Ala	Trp	Ile	Val	Thr	Gly	Gly	Leu	His	Thr
	145				150				155					160	
Gly	Ile	Gly	Arg	His	Val	Gly	Val	Ala	Val	Arg	Asp	His	Gln	Met	Ala
				165					170					175	
Ser	Thr	Gly	Gly	Thr	Lys	Val	Val	Ala	Met	Gly	Val	Ala	Pro	Trp	Gly
			180					185					190		
Val	Val	Arg	Asn	Arg	Asp	Thr	Leu	Ile	Asn	Pro	Lys	Gly	Ser	Phe	Pro
	195					200						205			
Ala	Arg	Tyr	Arg	Trp	Arg	Gly	Asp	Pro	Glu	Asp	Gly	Val	Gln	Phe	Pro
	210				215						220				
Leu	Asp	Tyr	Asn	Tyr	Ser	Ala	Phe	Phe	Leu	Val	Asp	Asp	Gly	Thr	His
	225				230				235					240	

-47-

Gly	Cys	Leu	Gly	Gly	Glu	Asn	Arg	Phe	Arg	Leu	Arg	Leu	Glu	Ser	Tyr
				245					250					255	
Ile	Ser	Gln	Gln	Lys	Thr	Gly	Val	Gly	Gly	Thr	Gly	Ile	Asp	Ile	Pro
			260					265					270		
Val	Leu	Leu	Leu	Leu	Ile	Asp	Gly	Asp	Glu	Lys	Met	Leu	Thr	Arg	Ile
		275					280					285			
Glu	Asn	Ala	Thr	Gln	Ala	Gln	Leu	Pro	Cys	Leu	Leu	Val	Ala	Gly	Ser
	290					295				300					
Gly	Gly	Ala	Ala	Asp	Cys	Leu	Ala	Glu	Thr	Leu	Glu	Asp	Thr	Leu	Ala
305				310					315						320
Pro	Gly	Ser	Gly	Gly	Ala	Arg	Gln	Gly	Glu	Ala	Arg	Asp	Arg	Ile	Arg
			325					330						335	
Arg	Phe	Phe	Pro	Lys	Gly	Asp	Leu	Glu	Val	Leu	Gln	Ala	Gln	Val	Glu
			340				345						350		
Arg	Ile	Met	Thr	Arg	Lys	Glu	Leu	Leu	Thr	Val	Tyr	Ser	Ser	Glu	Asp
	355					360						365			
Gly	Ser	Glu	Glu	Phe	Glu	Thr	Ile	Val	Leu	Lys	Ala	Leu	Val	Lys	Ala
	370					375					380				
Cys	Gly	Ser	Ser	Glu	Ala	Ser	Ala	Tyr	Leu	Asp	Glu	Leu	Arg	Leu	Ala
385				390					395						400
Val	Ala	Trp	Asn	Arg	Val	Asp	Ile	Ala	Gln	Ser	Glu	Leu	Phe	Arg	Gly
			405					410						415	
Asp	Ile	Gln	Trp	Arg	Ser	Phe	His	Leu	Glu	Ala	Ser	Leu	Met	Asp	Ala
		420					425					430			
Leu	Leu	Asn	Asp	Arg	Pro	Glu	Phe	Val	Arg	Leu	Leu	Ile	Ser	His	Gly
	435					440						445			
Leu	Ser	Leu	Gly	His	Phe	Leu	Thr	Pro	Met	Arg	Leu	Ala	Gln	Leu	Tyr
	450				455					460					
Ser	Ala	Ala	Pro	Ser	Asn	Ser	Leu	Ile	Arg	Asn	Leu	Leu	Asp	Gln	Ala
	465				470					475					480
Ser	His	Ser	Ala	Gly	Thr	Lys	Ala	Pro	Ala	Leu	Lys	Gly	Gly	Ala	Ala
			485					490						495	
Glu	Leu	Arg	Pro	Pro	Asp	Val	Gly	His	Val	Leu	Arg	Met	Leu	Leu	Gly
			500				505						510		
Lys	Met	Cys	Ala	Pro	Arg	Tyr	Pro	Ser	Gly	Gly	Ala	Trp	Asp	Pro	His
	515						520					525			
Pro	Gly	Gln	Gly	Phe	Gly	Glu	Ser	Met	Tyr	Leu	Leu	Ser	Asp	Lys	Ala
	530					535				540					
Thr	Ser	Pro	Leu	Ser	Leu	Asp	Ala	Gly	Leu	Gly	Gln	Ala	Pro	Trp	Ser
	545				550				555						560
Asp	Leu	Leu	Leu	Trp	Ala	Leu	Leu	Leu	Asn	Arg	Ala	Gln	Met	Ala	Met
			565					570						575	
Tyr	Phe	Trp	Glu	Met	Gly	Ser	Asn	Ala	Val	Ser	Ser	Ala	Leu	Gly	Ala
			580				585						590		
Cys	Leu	Leu	Leu	Arg	Val	Met	Ala	Arg	Leu	Glu	Pro	Asp	Ala	Glu	Glu
	595					600						605			
Ala	Ala	Arg	Arg	Lys	Asp	Leu	Ala	Phe	Lys	Phe	Glu	Gly	Met	Gly	Val
	610					615				620					
Asp	Leu	Phe	Gly	Glu	Cys	Tyr	Arg	Ser	Ser	Glu	Val	Arg	Ala	Ala	Arg
	625				630					635					640
Leu	Leu	Leu	Arg	Arg	Cys	Pro	Leu	Trp	Gly	Asp	Ala	Thr	Cys	Leu	Gln
			645						650					655	
Leu	Ala	Met	Gln	Ala	Asp	Ala	Arg	Ala	Phe	Phe	Ala	Gln	Asp	Gly	Val
			660				665						670		
Gln	Ser	Leu	Leu	Thr	Gln	Lys	Trp	Trp	Gly	Asp	Met	Ala	Ser	Thr	Thr
	675						680					685			
Pro	Ile	Trp	Ala	Leu	Val	Leu	Ala	Phe	Phe	Cys	Pro	Pro	Leu	Ile	Tyr
	690					695					700				
Thr	Arg	Leu	Ile	Thr	Phe	Arg	Lys	Ser	Glu	Glu	Glu	Pro	Thr	Arg	Glu
	705				710					715					720

-48-

Glu	Leu	Glu	Phe	Asp	Met	Asp	Ser	Val	Ile	Asn	Gly	Glu	Gly	Pro	Val
				725					730					735	
Gly	Thr	Ala	Asp	Pro	Ala	Glu	Lys	Thr	Pro	Leu	Gly	Val	Pro	Arg	Gln
			740					745					750		
Ser	Gly	Arg	Pro	Gly	Cys	Cys	Gly	Gly	Arg	Cys	Gly	Gly	Arg	Arg	Cys
		755					760					765			
Leu	Arg	Arg	Trp	Phe	His	Phe	Trp	Gly	Ala	Pro	Val	Thr	Ile	Phe	Met
	770					775					780				
Gly	Asn	Val	Val	Ser	Tyr	Leu	Leu	Phe	Leu	Leu	Leu	Phe	Ser	Arg	Val
785					790					795					800
Leu	Leu	Val	Asp	Phe	Gln	Pro	Ala	Pro	Pro	Gly	Ser	Leu	Glu	Leu	Leu
				805					810						815
Leu	Tyr	Phe	Trp	Ala	Phe	Thr	Leu	Leu	Cys	Glu	Glu	Leu	Arg	Gln	Gly
			820					825					830		
Leu	Ser	Gly	Gly	Gly	Gly	Ser	Leu	Ala	Ser	Gly	Gly	Pro	Gly	Pro	Gly
		835					840					845			
His	Ala	Ser	Leu	Ser	Gln	Arg	Leu	Arg	Leu	Tyr	Leu	Ala	Asp	Ser	Trp
	850					855					860				
Asn	Gln	Cys	Asp	Leu	Val	Ala	Leu	Thr	Cys	Phe	Leu	Leu	Gly	Val	Gly
865					870					875					880
Cys	Arg	Leu	Thr	Pro	Gly	Leu	Tyr	His	Leu	Gly	Arg	Thr	Val	Leu	Cys
				885				890						895	
Ile	Asp	Phe	Met	Val	Phe	Thr	Val	Arg	Leu	Leu	His	Ile	Phe	Thr	Val
			900					905					910		
Asn	Lys	Gln	Leu	Gly	Pro	Lys	Ile	Val	Ile	Val	Ser	Lys	Met	Met	Lys
		915					920					925			
Asp	Val	Phe	Phe	Phe	Leu	Phe	Phe	Leu	Gly	Val	Trp	Leu	Val	Ala	Tyr
	930					935				940					
Gly	Val	Ala	Thr	Glu	Gly	Leu	Leu	Arg	Pro	Arg	Asp	Ser	Asp	Phe	Pro
945					950					955					960
Ser	Ile	Leu	Arg	Arg	Val	Phe	Tyr	Arg	Pro	Tyr	Leu	Gln	Ile	Phe	Gly
				965				970						975	
Gln	Ile	Pro	Gln	Glu	Asp	Met	Asp	Val	Ala	Leu	Met	Glu	His	Ser	Asn
			980					985					990		
Cys	Ser	Ser	Glu	Pro	Gly	Phe	Trp	Ala	His	Pro	Pro	Gly	Ala	Gln	Ala
		995					1000					1005			
Gly	Thr	Cys	Val	Ser	Gln	Tyr	Ala	Asn	Trp	Leu	Val	Val	Leu	Leu	Leu
	1010					1015					1020				
Val	Ile	Phe	Leu	Leu	Val	Ala	Asn	Ile	Leu	Leu	Val	Asn	Leu	Leu	Ile
1025					1030					1035					1040
Ala	Met	Phe	Ser	Tyr	Thr	Phe	Gly	Lys	Val	Gln	Gly	Asn	Ser	Asp	Leu
				1045				1050						1055	
Tyr	Trp	Lys	Ala	Gln	Arg	Tyr	Arg	Leu	Ile	Arg	Glu	Phe	His	Ser	Arg
			1060					1065					1070		
Pro	Ala	Leu	Ala	Pro	Pro	Phe	Ile	Val	Ile	Ser	His	Leu	Arg	Leu	Leu
		1075				1080					1085				
Leu	Arg	Gln	Leu	Cys	Arg	Arg	Pro	Arg	Ser	Pro	Gln	Pro	Ser	Ser	Pro
	1090					1095					1100				
Ala	Leu	Glu	His	Phe	Arg	Val	Tyr	Leu	Ser	Lys	Glu	Ala	Glu	Arg	Lys
1105					1110					1115					1120
Leu	Leu	Thr	Trp	Glu	Ser	Val	His	Lys	Glu	Asn	Phe	Leu	Leu	Ala	Arg
				1125					1130					1135	
Ala	Arg	Asp	Lys	Arg	Glu	Ser	Asp	Ser	Glu	Arg	Leu	Lys	Arg	Thr	Ser
			1140					1145					1150		
Gln	Lys	Val	Asp	Leu	Ala	Leu	Lys	Gln	Leu	Gly	His	Ile	Arg	Glu	Tyr
		1155					1160					1165			
Glu	Gln	Arg	Leu	Lys	Val	Leu	Glu	Arg	Glu	Val	Gln	Gln	Cys	Ser	Arg
	1170					1175					1180				
Val	Leu	Gly	Trp	Val	Ala	Glu	Ala	Leu	Ser	Arg	Ser	Ala	Leu	Leu	Pro
1185					1190					1195					1200

Pro Gly Gly Pro Pro Pro Pro Asp Leu Pro Gly Ser Lys Asp
1205 1210

<210> 31
<211> 4646
<212> DNA
<213> Homo Sapiens

<400> 31
tcgacccacg cgtccgccca cgcgtccgcc cagcgcgtccg cccacgcgtc cgccccacgcg 60
tccgcccacg cgtccgggggt gaaagmramy cmygcktsms aaaaaccgtc acttaggaaa 120
agatgtcctt tcgggcagcc aggtcagca tgaggacacag aaggaaatgac actctggaca 180
gcacccggac cctgtactcc agcgcgtctc ggagcacaga cttgtcttac agtgaaagcg 240
acttggtgaa ttttattcaa gcaaatTTta agaaacgaga atgtgtcttc tttaccaaag 300
attccaaggg cagggagaat gtgtgcaagt gtggctatgc ccagagccag cacatggaag 360
gcacccagat caaccaaagt gagaaatgga actacaagaa acacaccaag gaatttccta 420
ccgacgcctt tggggatatt cagtttgaga cactggggaa gaaagggaag tatatacgtc 480
tgtctgcga caccgacgcg gaaatccttt acgagctgct gaccacgac tggcacctga 540
aaacacccaa cctggtcatt tctgtgaccg ggggcgcgca gaacttcgcc ctgaagccgc 600
gcatgcgcaa gatcttcagc cggctcatct acatcgcgca gtccaaaggt gcttggttc 660
tcacggggagg caccattat ggctgatga agtacatcg ggaggtggtg agagataaca 720
ccatcagcag gacttcagag gagaatattg tgccattgg catagcagct tggggcatgg 780
tctccaaccg ggacaccctc atcaggaatt gcgatgctga gggctatttt ttagccagct 840
accttatgga tgacttcaca agagatccac tgtgtatcct ggacaacaac cacacacatt 900
tgctgctcgt ggacaatggc tgtcatggac atcccactgt cgaagcaaaag ctccggaatc 960
agctagagaa gtatatctct gagcgacta ttcaagattc caactatggt ggcaagatcc 1020
ccattgtgtg ttttgccaa ggaggtggaa aagagacttt gaaagccatc aatacctcca 1080
tcaaaaataa aattccttgt gtgggtggtg aaggctcggg ccagatcgct gatgtgatcg 1140
ctagcctggt ggaggtggag gatgccctga catcttctgc cgtcaaggag aagctggtgc 1200
gctttttacc ccgcacggtg tcccgctgc ctgaggagga gactgagagt tggatcaaat 1260
ggctcaaaga aattctcgaa tgttctcacc tattaacagt tattaataat gaagaagctg 1320
gggatgaaat tgtgagcaat gccatctcct acgctctata caaagccttc agcaccagtg 1380
agcaagacaa ggataactgg aatgggcagc tgaagcttct gctggagtgg aaccagctgg 1440
acttagccaa tgatgagatt ttcaccaatg accgccgatg ggagtctgct gaccttcaag 1500
aagtcattgt tacggctctc ataaaggaca gacccaagt ttgtccgctc tttctggaga 1560
atggcttgaa cctacggaag tttctcacc atgatgtcct cactgaactc ttctccaacc 1620
acttcagcac gcttgtgtac cggaatctgc agatcgccaa gaattcctat aatgatgcc 1680
tcctcacgtt tgtctggaaa ctggttgca acttccgaag aggcctccgg aaggaaagaca 1740
gaaatggccg ggacgagatg gacatagaac tccacgacgt gtctcctatt actcggcacc 1800
ccctgcaagc tctcttcac tgggccattc ttcagaataa gaaggaaactc tccaaagtca 1860
tttgggagca gaccaggggc tgcactctgg cagccctggg agccagcaag cttctgaaga 1920
ctctggccaa agtgaagaac gacatcaatg ctgctgggga gtccgaggag ctggctaattg 1980
agtacgagac ccgggctgtt gagctgttca ctgagtgtta cagcagcgat gaagacttgg 2040
cagaacagct gctggtctat tctgtgaag cttgggggtg aagcaactgt ctggagctgg 2100
cggtggaagg cacagaccag catttcatcg cccagcctgg ggtccagaat tttctttcta 2160
agcaatggtg tggagagatt tcccgagaca ccaagaactg gaagattatc ctgtgtctgt 2220
ttattatacc cttggtgggc tgtggccttg tatcatttag gaagaaacct gtcgacaagc 2280
acaagaagct gctttggtac tatgtggcgt tcttcacctc tgctgtttgc ctacgtgctg ctcatggatt 2340
ggaatgtggt cttctacatc gccttctctc tctgtactc gctggtcttt gtccctctct 2400
tccattcgggt gccacacccc ccgagctgg tctgtactc gctggtcttt gtccctctct 2460
gtgatgaagt gagacagtgg tacgtaaatg ggggtgaatta ttttactgac ctgtggaatg 2520
tgatggacac gctggggcctt ttttacttca tagcaggaat tgtatttcgg ctccactctt 2580
ctaataaaaag ctctttgtat tctggacgag tcattttctg tctggactac attattttca 2640
ctctaagatt gatccacatt tttactgtaa gcagaaactt aggacccaag attataatgc 2700
tgcagaggat gctgatcgat gtgttcttct tctgttctct ctttgcgggtg tggatggtg 2760
cctttggcgt ggccaggcaa gggatcctta ggcagaatga gcagcgctgg aggtggatat 2820
tccgttcgggt catctacgag ccctacctgg ccctggtcgg ccaggtgccc agtgacgtgg 2880
atggtaccac gtatgacttt gcccaactgca ccttcaactgg gaatgagtc aagccactgt 2940
gtgtggagct ggatgagcac aacctgcccc ggttccccga gtggatcacc atccccctgg 3000
tgtgcatcta catgttatcc accaaccatcc tgcgtgtgcaa cctgctggtc gccatggttg 3060

-50-

gctacacggt	gggcaccgctc	caggagaaca	atgaccaggt	ctggaagttc	cagaggtact	3120
tcctgggtgca	ggagtactgc	agccgcctca	atatccccctt	cccccttcac	gtcttcgctt	3180
acttctacat	ggtggtgaag	aagtgccttca	agtgttgctg	caaggagaaa	aacatggagt	3240
cttctgtctg	ctgtttcaaa	aatgaagaca	atgagactct	ggcatgggag	ggtgtcatga	3300
aggaaaacta	ccttgtcaag	atcaacacaa	aagccaacga	cacctcagag	gaaatgagggc	3360
atcgatttag	acaactggat	acaaagctta	atgatctcaa	gggtcttctg	aaagagattg	3420
ctaataaaat	caaataaaac	tgtatgaact	ctaattggaga	aaaatctaata	tatagcaaga	3480
tcatattaag	gaatgctgat	gaacaatttt	gctatcgact	actaaatgag	agattttcag	3540
accctgggt	acatgggtga	tgattttaaa	tcaccctagt	gtgctgagac	cttgagaata	3600
aagtgtgtga	ttggtttcat	acttgaagac	ggatataaag	gaagaatatt	tcctttatgt	3660
gtttctccag	aatgggtgcct	gtttctctct	gtgtctcaat	gcctgggact	ggaggttgat	3720
agtttaagt	tggtcttacc	gcctcctttt	tcctttaatc	ttatttttga	tgaacacata	3780
tataggagaa	catctatcct	atgaataaga	acctggtcac	gctttactcc	tgtattgtta	3840
ttttgttcat	ttccaattga	ttctctactt	ttcccttttt	tgtattatgt	gactaattag	3900
ttggcatatt	gtwaaaagtc	tctcaaatta	ggccagattc	taaaacatgc	tgacagcaaga	3960
ggacccccgt	ctcttcagga	aaagtgtttt	catttctcag	gatgcttctt	acctgtcaga	4020
ggaggtgaca	aggcagtcct	ttgctctctt	ggactcacca	ggctcctatt	gaaggaaacca	4080
ccccattcc	taaatatgtg	aaaagtcgcc	caaaatgcaa	ccttgaaagg	cactactgac	4140
tttgttctta	ttggatactc	ctcttattta	ttatttttcc	attaaaaata	atagctggct	4200
attatagaaa	atttagacca	tacagagatg	tagaaagaac	ataaattgtc	cccattacct	4260
taaggtaatc	actgctaaca	atctctggat	ggtttttcaa	gtctattttt	tttctatgta	4320
tgtctcaatt	ctctttcaaa	attttacaga	atgttatcat	actacatata	tactttttat	4380
gtaagctttt	tcacttagta	ttttatcaaa	tatgttttta	ttatatcat	agccttctta	4440
aacattatat	caataattgc	ataataggca	acctctagcg	attaccataa	ttttgctcat	4500
tgaaggctat	ctccagttga	tcattgggat	gagcatcttt	gtgcatgaat	cctattgctg	4560
tatttgggaa	aattttccaa	ggttagattc	caataaatat	ctatttatta	ttaaaaaaa	4620
aaaaaaaag	gcggccgctc	tagagt				4646

<210> 32
 <211> 1104
 <212> PRT
 <213> Homo Sapiens

<400> 32

Met	Ser	Phe	Arg	Ala	Ala	Arg	Leu	Ser	Met	Arg	Asn	Arg	Arg	Asn	Asp
1				5					10					15	
Thr	Leu	Asp	Ser	Thr	Arg	Thr	Leu	Tyr	Ser	Ser	Ala	Ser	Arg	Ser	Thr
			20					25					30		
Asp	Leu	Ser	Tyr	Ser	Glu	Ser	Asp	Leu	Val	Asn	Phe	Ile	Gln	Ala	Asn
			35				40					45			
Phe	Lys	Lys	Arg	Glu	Cys	Val	Phe	Phe	Thr	Lys	Asp	Ser	Lys	Ala	Thr
	50				55						60				
Glu	Asn	Val	Cys	Lys	Cys	Gly	Tyr	Ala	Gln	Ser	Gln	His	Met	Glu	Gly
	65				70				75					80	
Thr	Gln	Ile	Asn	Gln	Ser	Glu	Lys	Trp	Asn	Tyr	Lys	Lys	His	Thr	Lys
			85					90					95		
Glu	Phe	Pro	Thr	Asp	Ala	Phe	Gly	Asp	Ile	Gln	Phe	Glu	Thr	Leu	Gly
			100					105					110		
Lys	Lys	Gly	Lys	Tyr	Ile	Arg	Leu	Ser	Cys	Asp	Thr	Asp	Ala	Glu	Ile
		115					120					125			
Leu	Tyr	Glu	Leu	Leu	Thr	Gln	His	Trp	His	Leu	Lys	Thr	Pro	Asn	Leu
	130					135					140				
Val	Ile	Ser	Val	Thr	Gly	Gly	Ala	Lys	Asn	Phe	Ala	Leu	Lys	Pro	Arg
	145				150					155				160	
Met	Arg	Lys	Ile	Phe	Ser	Arg	Leu	Ile	Tyr	Ile	Ala	Gln	Ser	Lys	Gly
				165					170					175	
Ala	Trp	Ile	Leu	Thr	Gly	Gly	Thr	His	Tyr	Gly	Leu	Met	Lys	Tyr	Ile
			180					185					190		
Gly	Glu	Val	Val	Arg	Asp	Asn	Thr	Ile	Ser	Arg	Ser	Ser	Glu	Glu	Asn
		195					200					205			

-51-

Ile	Val	Ala	Ile	Gly	Ile	Ala	Ala	Trp	Gly	Met	Val	Ser	Asn	Arg	Asp
210						215					220				
Thr	Leu	Ile	Arg	Asn	Cys	Asp	Ala	Glu	Gly	Tyr	Phe	Leu	Ala	Gln	Tyr
225				230						235					240
Leu	Met	Asp	Asp	Phe	Thr	Arg	Asp	Pro	Leu	Cys	Ile	Leu	Asp	Asn	Asn
				245					250					255	
His	Thr	His	Leu	Leu	Val	Asp	Asn	Gly	Cys	His	Gly	His	Pro	Thr	
			260					265				270			
Val	Glu	Ala	Lys	Leu	Arg	Asn	Gln	Leu	Glu	Lys	Tyr	Ile	Ser	Glu	Arg
		275					280					285			
Thr	Ile	Gln	Asp	Ser	Asn	Tyr	Gly	Gly	Lys	Ile	Pro	Ile	Val	Cys	Phe
290					295						300				
Ala	Gln	Gly	Gly	Gly	Lys	Glu	Thr	Leu	Lys	Ala	Ile	Asn	Thr	Ser	Ile
305					310					315					320
Lys	Asn	Lys	Ile	Pro	Cys	Val	Val	Val	Glu	Gly	Ser	Gly	Gln	Ile	Ala
				325					330					335	
Asp	Val	Ile	Ala	Ser	Leu	Val	Glu	Val	Glu	Asp	Ala	Leu	Thr	Ser	Ser
			340					345					350		
Ala	Val	Lys	Glu	Lys	Leu	Val	Arg	Phe	Leu	Pro	Arg	Thr	Val	Ser	Arg
		355					360					365			
Leu	Pro	Glu	Glu	Glu	Thr	Glu	Ser	Trp	Ile	Lys	Trp	Leu	Lys	Glu	Ile
370						375					380				
Leu	Glu	Cys	Ser	His	Leu	Leu	Thr	Val	Ile	Lys	Met	Glu	Glu	Ala	Gly
385					390					395					400
Asp	Glu	Ile	Val	Ser	Asn	Ala	Ile	Ser	Tyr	Ala	Leu	Tyr	Lys	Ala	Phe
				405					410					415	
Ser	Thr	Ser	Glu	Gln	Asp	Lys	Asp	Asn	Trp	Asn	Gly	Gln	Leu	Lys	Leu
			420					425					430		
Leu	Leu	Glu	Trp	Asn	Gln	Leu	Asp	Leu	Ala	Asn	Asp	Glu	Ile	Phe	Thr
		435					440					445			
Asn	Asp	Arg	Arg	Trp	Glu	Ser	Ala	Asp	Leu	Gln	Glu	Val	Met	Phe	Thr
		450				455					460				
Ala	Leu	Ile	Lys	Asp	Arg	Pro	Lys	Phe	Val	Arg	Leu	Phe	Leu	Glu	Asn
465					470					475					480
Gly	Leu	Asn	Leu	Arg	Lys	Phe	Leu	Thr	His	Asp	Val	Leu	Thr	Glu	Leu
				485					490					495	
Phe	Ser	Asn	His	Phe	Ser	Thr	Leu	Val	Tyr	Arg	Asn	Leu	Gln	Ile	Ala
			500					505					510		
Lys	Asn	Ser	Tyr	Asn	Asp	Ala	Leu	Leu	Thr	Phe	Val	Trp	Lys	Leu	Val
		515					520					525			
Ala	Asn	Phe	Arg	Arg	Gly	Phe	Arg	Lys	Glu	Asp	Arg	Asn	Gly	Arg	Asp
		530				535					540				
Glu	Met	Asp	Ile	Glu	Leu	His	Asp	Val	Ser	Pro	Ile	Thr	Arg	His	Pro
545					550					555					560
Leu	Gln	Ala	Leu	Phe	Ile	Trp	Ala	Ile	Leu	Gln	Asn	Lys	Lys	Glu	Leu
				565					570					575	
Ser	Lys	Val	Ile	Trp	Glu	Gln	Thr	Arg	Gly	Cys	Thr	Leu	Ala	Ala	Leu
			580					585					590		
Gly	Ala	Ser	Lys	Leu	Leu	Lys	Thr	Leu	Ala	Lys	Val	Lys	Asn	Asp	Ile
		595					600					605			
Asn	Ala	Ala	Gly	Glu	Ser	Glu	Glu	Leu	Ala	Asn	Glu	Tyr	Glu	Thr	Arg
		610				615					620				
Ala	Val	Glu	Leu	Phe	Thr	Glu	Cys	Tyr	Ser	Ser	Asp	Glu	Asp	Leu	Ala
625					630					635					640
Glu	Gln	Leu	Leu	Val	Tyr	Ser	Cys	Glu	Ala	Trp	Gly	Gly	Ser	Asn	Cys
				645					650					655	
Leu	Glu	Leu	Ala	Val	Glu	Ala	Thr	Asp	Gln	His	Phe	Ile	Ala	Gln	Pro
			660					665					670		
Gly	Val	Gln	Asn	Phe	Leu	Ser	Lys	Gln	Trp	Tyr	Gly	Glu	Ile	Ser	Arg
		675					680					685			

-52-

Asp	Thr	Lys	Asn	Trp	Lys	Ile	Ile	Leu	Cys	Leu	Phe	Ile	Ile	Pro	Leu
690						695					700				
Val	Gly	Cys	Gly	Phe	Val	Ser	Phe	Arg	Lys	Lys	Pro	Val	Asp	Lys	His
705					710					715					720
Lys	Lys	Leu	Leu	Trp	Tyr	Tyr	Val	Ala	Phe	Phe	Thr	Ser	Pro	Phe	Val
				725					730					735	
Val	Phe	Ser	Trp	Asn	Val	Val	Phe	Tyr	Ile	Ala	Phe	Leu	Leu	Leu	Phe
			740					745					750		
Ala	Tyr	Val	Leu	Leu	Met	Asp	Phe	His	Ser	Val	Pro	His	Pro	Pro	Glu
		755					760					765			
Leu	Val	Leu	Tyr	Ser	Leu	Val	Phe	Val	Leu	Phe	Cys	Asp	Glu	Val	Arg
	770					775					780				
Gln	Trp	Tyr	Val	Asn	Gly	Val	Asn	Tyr	Phe	Thr	Asp	Leu	Trp	Asn	Val
785					790					795					800
Met	Asp	Thr	Leu	Gly	Leu	Phe	Tyr	Phe	Ile	Ala	Gly	Ile	Val	Phe	Arg
				805					810					815	
Leu	His	Ser	Ser	Asn	Lys	Ser	Ser	Leu	Tyr	Ser	Gly	Arg	Val	Ile	Phe
				820				825					830		
Cys	Leu	Asp	Tyr	Ile	Ile	Phe	Thr	Leu	Arg	Leu	Ile	His	Ile	Phe	Thr
		835					840					845			
Val	Ser	Arg	Asn	Leu	Gly	Pro	Lys	Ile	Ile	Met	Leu	Gln	Arg	Met	Leu
	850					855					860				
Ile	Asp	Val	Phe	Phe	Phe	Leu	Phe	Leu	Phe	Ala	Val	Trp	Met	Val	Ala
865					870					875					880
Phe	Gly	Val	Ala	Arg	Gln	Gly	Ile	Leu	Arg	Gln	Asn	Glu	Gln	Arg	Trp
				885					890					895	
Arg	Trp	Ile	Phe	Arg	Ser	Val	Ile	Tyr	Glu	Pro	Tyr	Leu	Ala	Met	Phe
			900					905					910		
Gly	Gln	Val	Pro	Ser	Asp	Val	Asp	Gly	Thr	Thr	Tyr	Asp	Phe	Ala	His
		915					920					925			
Cys	Thr	Phe	Thr	Gly	Asn	Glu	Ser	Lys	Pro	Leu	Cys	Val	Glu	Leu	Asp
	930					935					940				
Glu	His	Asn	Leu	Pro	Arg	Phe	Pro	Glu	Trp	Ile	Thr	Ile	Pro	Leu	Val
945					950					955					960
Cys	Ile	Tyr	Met	Leu	Ser	Thr	Asn	Ile	Leu	Leu	Val	Asn	Leu	Leu	Val
				965					970					975	
Ala	Met	Phe	Gly	Tyr	Thr	Val	Gly	Thr	Val	Gln	Glu	Asn	Asn	Asp	Gln
			980					985					990		
Val	Trp	Lys	Phe	Gln	Arg	Tyr	Phe	Leu	Val	Gln	Glu	Tyr	Cys	Ser	Arg
		995					1000					1005			
Leu	Asn	Ile	Pro	Phe	Pro	Phe	Ile	Val	Phe	Ala	Tyr	Phe	Tyr	Met	Val
	1010					1015					1020				
Val	Lys	Lys	Cys	Phe	Lys	Cys	Cys	Cys	Lys	Glu	Lys	Asn	Met	Glu	Ser
1025					1030					1035					104
Ser	Val	Cys	Cys	Phe	Lys	Asn	Glu	Asp	Asn	Glu	Thr	Leu	Ala	Trp	Glu
				1045					1050					1055	
Gly	Val	Met	Lys	Glu	Asn	Tyr	Leu	Val	Lys	Ile	Asn	Thr	Lys	Ala	Asn
			1060					1065					1070		
Asp	Thr	Ser	Glu	Glu	Met	Arg	His	Arg	Phe	Arg	Gln	Leu	Asp	Thr	Lys
		1075					1080					1085			
Leu	Asn	Asp	Leu	Lys	Gly	Leu	Leu	Lys	Glu	Ile	Ala	Asn	Lys	Ile	Lys
	1090					1095						1100			

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/29996

A. CLASSIFICATION OF SUBJECT MATTER IPC 7 C07K14/705 C12N15/12 C12Q1/68 C12N5/10 C07K16/28 G01N33/53 A61K38/17		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED Minimum documentation searched (classification system followed by classification symbols) IPC 7 C12N C07K C12Q A61K G01N		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practical, search terms used) BIOSIS, EPO-Internal, WPI Data, PAJ, MEDLINE, SCISEARCH, EMBASE, BIOTECHNOLOGY ABS, CHEM ABS Data, STRAND, GENSEQ, EMBL		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	DATABASE GENEMBL 'Online! 16 February 1998 (1998-02-16) STRAUSBERG, R.: "ob70f05.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1336737 3', mRNA sequence" XP002138823 Accession AA809355	1,2, 6-19, 25-35
X	DATABASE GENEMBL 'Online! 10 July 1998 (1998-07-10) MARRA ET AL.: "ub28d10.r1 Soares 2NbMT Mus musculus cDNA clone IMAGE:1379059 5' mRNA sequence" XP002149803 Accession AI050262	1,6-19, 25-35
<div style="text-align: center;">-/-</div>		
<input checked="" type="checkbox"/> Further documents are listed in the continuation of box C. <input checked="" type="checkbox"/> Patent family members are listed in annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier document but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art. "&" document member of the same patent family		
Date of the actual completion of the international search 16 October 2000		Date of mailing of the international search report 30. 10. 00
Name and mailing address of the ISA European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016		Authorized officer ALCONADA RODRIG., A

INTERNATIONAL SEARCH REPORT

International Application No

PCT/US 99/29996

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	<p>DATABASE GENEMBL 'Online! 19 July 1997 (1997-07-19) STRAUSBERG, R.: "nt64e11.s1 NCI_CGAP_Pr12 Homo sapiens cDNA clone IMAGE:981644 mRNA sequence" XP002148641 Accession AA523749</p>	1,3, 10-19, 25-35
X	<p>WO 98 15657 A (ABBOTT LAB) 16 April 1998 (1998-04-16)</p>	1,4, 6-19, 25-35
Y	<p>page 4, line 7 -page 5, line 13 page 5, line 24 -page 7, line 28 SEQ ID NOs. 9 and 25</p>	20-24
X	<p>WO 98 37093 A (CORIXA CORP) 27 August 1998 (1998-08-27)</p>	1,4, 6-19, 25-35
Y	<p>page 7, paragraph 2 page 9, paragraphs 2,3 page 13 -page 17 page 21, paragraph 3 SEQ ID NOs: 109 and 112</p>	20-24
X	<p>DATABASE GENEMBL 'Online! 18 November 1997 (1997-11-18) STRAUSBERG, R.: "nt76b07.s1 NCI_CGAP_Pr3 Homo sapiens cDNA clone IMAGE:1204405, mRNA" XP002148642 Accession AA654650</p>	1,5-19, 25-35
Y	<p>Accession AA654650</p>	20-24
Y	<p>DATABASE GENEMBL 'Online! 30 November 1998 (1998-11-30) SHIMIZU, N.: "Homo sapiens mRNA complete cds." XP002148643 Accession number AB001535 -& NAGAMINE ET AL.: "Molecular cloning of a novel putative Ca²⁺ channel protein (TRPC7) highly expressed in brain" GENOMICS, vol. 54, 15 November 1998 (1998-11-15), pages 124-131, XP000938744 the whole document</p>	20-24

-/--

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/29996

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ZHU, XI ET AL: "Molecular cloning of a widely expressed human homologue for the <i>Drosophila trp</i> gene." FEBS LETTERS, (1995) VOL. 373, NO. 3, PP. 193-198., XP000907241 page 194; figures 1,3	20,21, 23,25, 26,28, 29,31
A	HUNTER JOHN J ET AL: "Chromosomal localization and genomic characterization of the mouse melastatin gene (<i>Mln1</i>)." GENOMICS NOV. 15, 1998, vol. 54, no. 1, 15 November 1998 (1998-11-15), pages 116-123, XP000910696 ISSN: 0888-7543 cited in the application page 119; figure 2	20,21,23
A	WES PAUL D ET AL: "TRPC1, a human homolog of a <i>Drosophila</i> store-operated channel." PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA 1995, vol. 92, no. 21, 1995, pages 9652-9656, XP002138820 ISSN: 0027-8424 the whole document	20,21, 23,25, 26,28, 29,31
A	ZHU, XI ET AL: "Trp, A novel mammalian gene family essential for agonist-activated capacitative Ca-2+ entry." CELL, vol. 85, no. 5, 1996, pages 661-671, XP000907242 page 662 page 665 figures 1,5,6	20,21, 25,26, 28,29,31
A	GARCIA REYNALDO L ET AL: "Differential expression of mammalian TRP homologues across tissues and cell lines." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS 1997, vol. 239, no. 1, 1997, pages 279-283, XP002138822 ISSN: 0006-291X See Materials and Methods figure 1	25,26, 28-30

INTERNATIONAL SEARCH REPORT

In ternational Application No

PCT/US 99/29996

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT		
Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	SINKINS WILLIAM G ET AL: "Functional expression of TrpC1: A human homologue of the Drosophila Trp channel." BIOCHEMICAL JOURNAL APRIL, 1998, vol. 331, no. 1, April 1998 (1998-04), pages 331-339, XP000864583 ISSN: 0264-6021 page 333-335; figures 3-5	24
A	PREUSS KLAUS-DIETER ET AL: "Expression and characterization of a trp1 homolog from rat." BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS NOV. 7, 1997, vol. 240, no. 1, 7 November 1997 (1997-11-07), pages 167-172, XP002138821 ISSN: 0006-291X figure 2	24
A	OBUKHOV, ALEXANDER G. ET AL: "Direct activation of trp1 cation channels by G-alpha-11 subunits." EMBO (EUROPEAN MOLECULAR BIOLOGY ORGANIZATION) JOURNAL, (1996) VOL. 15, NO. 21, PP. 5833-5838., XP000907243 figure 2	24
P,X	WO 99 09199 A (RYAZANOV ALEXEY G ;PAVUR KAREN S (US); HAIT WILLIAM N (US); UNIV M) 25 February 1999 (1999-02-25) see melanome kinase polynucleotide and polypeptide sequences on page 16-17	1,3, 10-19, 25-36
P,X	WO 99 09166 A (SHAPERO MICHAEL H ;DENDREON CORP (US); LAUS REINER (US); TSAVALER) 25 February 1999 (1999-02-25) page 17, line 24 -page 18, line 9 page 25, line 19-32 page 28, line 1-4 SEQ ID NOs: 27, 28 and 31.	1,5-19, 25-35
T	SCHARENBERG A M ET AL: "MLSN-1/SOC-1 defines a widely expressed Ca2+/cation channel family involved in Ca2+ homeostasis and store-operated Ca2+ signaling." FIFTY-THIRD ANNUAL MEETING OF THE SOCIETY OF GENERAL PHYSIOLOGISTS;WOODS HOLE, MASSACHUSETTS, USA; SEPTEMBER 9-11, 1999, vol. 114, no. 1, July 1999 (1999-07), page 14a XP000910708 Journal of General Physiology July, 1999 ISSN: 0022-1295	

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/29996

Box I Observations where certain claims were found unsearchable (Continuation of Item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:
2. ☒ Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210
3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

As a result of the prior review under R. 40.2(e) PCT,
no additional fees are to be refunded.

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.
2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.
3. ☒ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:
1-36
4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

- ☒ The additional search fees were accompanied by the applicant's protest.
☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Present claims 1-5, 10-13, 16-19, 32-35 relate to an extremely large number of possible polynucleotides, polypeptides encoded by them, binding polypeptides, and kits and pharmaceutical compositions containing said polypeptides and polynucleotides. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the compounds claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to the polynucleotide of SEQ ID NOs: 1, 27, 29 and 31 and the corresponding polypeptide of SEQ ID NOs: 2, 28, 30 and 32.

Present claims 16 and 17 relate to an extremely large number of possible compounds, namely, a polypeptide that binds to the polypeptide of the invention. Support within the meaning of Article 6 PCT and/or disclosure within the meaning of Article 5 PCT is to be found, however, for only a very small proportion of the compounds claimed. In the present case, the claims so lack support, and the application so lacks disclosure, that a meaningful search over the whole of the claimed scope is impossible. Consequently, the search has been carried out for those parts of the claims which appear to be supported and disclosed, namely those parts relating to an antibody, antibody fragment, F(ab)2 fragment or a fragment including a CDR3 region selective for the polypeptides of the invention.

The applicant's attention is drawn to the fact that claims, or parts of claims, relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. Claims: 1, 6-36 (partially) and 2 (complete)

An isolated nucleic acid molecule comprising a nucleic acid molecule that hybridizes to a nucleic acid molecule of SEQ ID NO:1 and which code for a SOC/CRAC polypeptide, nucleic acid molecules that differ in codon sequence due to degeneracy of the genetic code and complement thereof, polynucleotides which are not identical to the SEQ ID or sequences of GenBank accession number of Table 1; expression vector, host cells; polypeptide encoded thereof (SEQ ID NO:2); polypeptides binding to the polypeptide of SEQ ID NO:2, including antibodies; kits comprising agents that selectively bind to the polynucleotide (SEQ ID NO:1) or polypeptide (SEQ ID NO:2) of the invention; pharmaceutical compositions containing the polynucleotide or polypeptides of the invention; a method for isolating the SOC/CRAC molecule having SOC/CRAC calcium channel activity comprising contacting a binding molecule that is SOC/CRAC nucleic acid or a SOC/CRAC binding polypeptide with a sample containing SOC/CRAC molecules allowing the formation of the complex, detecting the formation of the complex, isolating the SOC/CRAC molecule and determining whether the isolated SOC/CRAC molecule has SOC/CRAC calcium channel activity; a method for identifying agents useful in the modulation of SOC/CRAC calcium channel activity; a method to determine the level of SOC/CRAC expression in a subject, including expression of SOC/CRAC polypeptide or mRNA in a tissue or biological fluid sample using PCR, Northern blotting, and mono- and polyclonal antisera and a method for identifying agents useful in the modulation of the SOC/CRAC polypeptide kinase activity, comprising the use of aminoacids 999-1180 from SEQ ID NO:4 as a candidate kinase.

2. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:3 and to the encoded polypeptide of SEQ ID NO:4

3. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:5 and to the encoded polypeptide of SEQ ID NO:6

4. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:7 and to the encoded polypeptide of SEQ ID NO:8

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

5. Claims: 1,6-36 (partially) and 37 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:23 and to the encoded polypeptide of SEQ ID NO:24

6. Claims: 1,6-36 (partially)

As subject 1, but referred to the polynucleotide of SEQ ID NO:25 and to the encoded polypeptide of SEQ ID NO:26

7. Claims: 1,10-36 (partially) and 3 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:27 and to the encoded polypeptide of SEQ ID NO:28

8. Claims: 1,6-36 (partially) and 4 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:29 and to the encoded polypeptide of SEQ ID NO:30

9. Claims: 1,6-36 (partially) and 5 (complete)

As subject 1, but referred to the polynucleotide of SEQ ID NO:31 and to the encoded polypeptide of SEQ ID NO:32.

INTERNATIONAL SEARCH REPORT

Information on patent family members

In ternational Application No

PCT/US 99/29996

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9815657 A	16-04-1998	US 5919638 A EP 0954599 A US 6110675 A	06-07-1999 10-11-1999 29-08-2000
WO 9837093 A	27-08-1998	AU 6181898 A CN 1252837 T EP 1005546 A NO 994069 A PL 335348 A ZA 9801585 A	09-09-1998 10-05-2000 07-06-2000 22-10-1999 25-04-2000 04-09-1998
WO 9909199 A	25-02-1999	AU 9110098 A	08-03-1999
WO 9909166 A	25-02-1999	AU 9021898 A EP 1005549 A	08-03-1999 07-06-2000

This Page Blank (uspto)